AD_____

Award Number: DAMD17-94-J-4109

TITLE: Regional Breast Cancer Screening Network

PRINCIPAL INVESTIGATOR: Robert Greenberg, M.D.
Patricia Carney, Ph.D.

CONTRACTING ORGANIZATION: Dartmouth College
Hanover, New Hampshire 03755-1404

REPORT DATE: September 2000

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 074-0188*



| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | September 2000 | Final (1 Oct 94 - 31 Aug 00) |

**4. TITLE AND SUBTITLE**
Regional Breast Cancer Screening Network

**5. FUNDING NUMBERS**
DAMD17-94-J-4109

**6. AUTHOR(S)**
Robert Greenberg, M.D.
Patricia Carney, Ph.D.

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Dartmouth College
Hanover, New Hampshire 03755-1404

E-MAIL:
e.robert.greenberg@dartmouth.edu

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for public release; distribution unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 Words)*

The long-term objective of this Project was to improve the health of New Hampshire (NH) women by enhancing breast cancer screening and detection. To accomplish this, the New Hampshire Mammography Network (NHMN) implemented a comprehensive database tracking system, allowing us to follow the outcomes of women receiving mammography (either diagnostic or screening) and other breast procedures (e.g. ultrasound, biopsy or fine needle aspiration) over time. To date, we have 256,197 mammographic encounters in our database, representing 166,664 NH women. As of September 2000, 22,398 breast pathology reports exist in the pathology database. Of these, 7,871 are matched (52.4%) to women in the NHMN. Of the matched reports, 63% are benign, 6% are atypical, 0.5% are suspicious, (6.5%) are non-invasive, 20% are invasive, and 3% are unsatisfactory. We have now identified a total of 1,538 non-invasive breast cancer cases and 5,132 invasive breast cancer cases. We have used this database to conduct several analyses and special research projects, which are described in detail in the body of this report. Overall, the support provided by the US Department of Defense to establish the NH Mammography Network has led to a strong and productive research program in breast cancer surveillance.

**14. SUBJECT TERMS**
Breast Cancer, Mammography, Epidemiology, Screening, Population Registry, Diagnostic Accuracy

**15. NUMBER OF PAGES**
228

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | Unlimited |

TABLE OF CONTENTS

## INTRODUCTION

The long-term objective of this Project was to improve the health of New Hampshire (NH) women by enhancing breast cancer screening and detection. To accomplish this, the New Hampshire Mammography Network (NHMN) implemented a comprehensive database tracking system, allowing us to follow the outcomes of women receiving mammography (either diagnostic or screening) and other breast procedures (e.g. ultrasound, biopsy or fine needle aspiration) over time. All mammography facilities in the state were visited in 1994 and provided with materials and mechanisms to collect and furnish data to the central data repository (NHMN). These materials included Project manuals, instruments and materials to promote ongoing enrollment by NH women in the project. In 1997 and 1998, we established linkages among various state agencies to obtain pertinent data, including the NH State Cancer Registry and the NH death clearance tapes. Accomplishing these linkages allowed us to conduct data analysis on the test characteristics of mammography (sensitivity, specificity, positive predictive value, negative predictive value) in NH and to provide feedback reports to NH radiologists and facilities regarding their performance. We have been generating reports every six months for facilities and radiologists since February 1998. We have developed a report handling policy to delineate appropriate uses of NHMN registry reports, which is included in **Appendix A**.

New Hampshire (NH) is well suited to this type of population-based research due to its stable population with a blend of urban and rural communities and a relatively high level of literacy (82.2% of New Hampshire adults are high school graduates), which simplified interviewing and form completion. New Hampshire is also a relatively small state with an estimated population of 1,136,000 (1). Breast cancer is the leading cancer in NH women with over 800 cases per year, representing 33% of all female cancers (2). The mortality rate is 29 per 100,000, which is higher than the national rate of 27.3 per 100,000 (3). Women between the ages of 40 and 74 represent about 14% of the population of 160,000 (1). Data from the Centers for Disease Control 1991 NH Behavioral Risk Factor Survey found that 37% of women between the ages of 40-49 report that they have not had a mammogram within the past two years and 50% of women over age 50 report that they have not had a mammogram within the past year (4). Clearly, the development of a population-based mammography registry is an important contribution to understanding the problem of breast cancer in New Hampshire.

To date, we have 256,197 mammographic encounters in our database, representing 166,664 NH women. As of September 2000, 22,398 breast pathology reports exist in the pathology database. Of these, 7,871 are matched (52.4%) to women in the NHMN. Of the matched reports, 63% are benign, 6% are atypical, 0.5% are suspicious, (6.5%) are non-invasive, 20% are invasive, and 3% are unsatisfactory. We have now identified a total of 1,538 non-invasive breast cancer cases and 5,132 invasive breast cancer cases. **Appendix B** contains a publication that describes the characteristics of women in our registry and the exams that they have

received, including our estimates of the penetration of mammography in the population by age category.

We applied for and received funding from the Centers for Disease Control in January 1997 (Carney, PA-PI) to conduct studies on both the interpretive agreement among community-based pathologists in breast cancer diagnosis and the accuracy and reproducibility of ductal carcinoma in situ grading systems. These projects have led to three publications, which are included in **Appendix C** (see **Section 3 Expanded Use of the Infrastructure**, page 14).

We have also received funding for five additional studies. One from the American Cancer Society (Carney, PA-PI), which is just beginning its third and final year, to study characteristics of women age 50 and older who do and do not adhere to interval screening (within 24 months). Another from the National Institute for Nursing Research (Shannon Award)(Carney, PA-PI) to assess how women's risk and anxiety traits influence their screening behavior. This study is just finishing this fall (9/2000). A third study (Carney, PA-PI) has been funded by the National Cancer Institute (4/2000) to study the relationships among hormone replacement therapy use and breast cancer incidence, detection, prognostic characteristics and health-related quality of life. A fourth study (Carney, PA-PI), funded by the National Cancer Institute (4/2000) will provide ongoing funding for the NHMN while conducting several special projects on breast cancer surveillance. These projects include:

1) Developing risk prediction models for both invasive and non-invasive breast cancer;
2) Comparing actual risk, perceived risk, and anxiety traits in a population-based sample of unscreened NH women to screened women in the NHMN database;
3) Evaluating the influence of menstrual cycle phase on breast density and mammographic performance;
4) Determining whether benign breast biopsy characteristics (biopsy type, number of breast biopsies and biopsy outcome) are related to mammographic accuracy; and
5) Developing a longitudinal model of mammography/health states defined by screening compliance, mammography outcomes, follow-up and disease outcomes in individual subjects, and determine predictors for transitions between these states.

This final aim will include an evaluation of mammography-related predictors for all cause mortality and breast cancer mortality. Lastly, we received funding from the Agency for Research in Health Quality (9/00) (Elmore, J- PI, Carney, PA-Co-PI) to study factors associated with variability in mammographic interpretation. Study findings from completed studies will be described under **Key Research Accomplishments** (page 25) and the studies that are just beginning will be described under **Reportable Outcomes** (page 27).

Overall, the support provided by the US Department of Defense to establish the NH Mammography Network has led to a strong and productive research program in breast cancer surveillance. We have additionally been active in the National Cancer Institute funded Breast Cancer Surveillance Consortium by submitting data collected on mammographic encounters in New Hampshire, taking the lead in developing a policy and procedure manual to insure data integrity and confidentiality at each Consortium site, and participating in the analysis and development of several manuscripts on analytic methods and findings from the pooled analyses. These include: The Effect of Changing Definitions on Performance Indices for Mammography (In Review), Use of BI-RADS in Screening Mammography (In Review), Use of BI-RADS in Diagnostic Mammography (In Development), The Performance of Mammography Among Women with and without a First Degree Relative with Breast Cancer (In Press). The policy and procedure manual on data integrity confidentiality is included in **Appendix D**. Two recent publications that describe the Consortium and the medico-legal analysis we conducted to insure legal protection of the data at Consortium member sites and the Statistical Coordinating Center (to which all data are sent for pooled analysis) are included in **Appendix E**.

We will address in the Body of this report the progress we made in accomplishing the tasks we outlined in our original proposal. We will outline these in three sections: Establishing Data Collection Procedures; Data Analysis and Feedback Reporting Procedures; and Expanded Use of the Infrastructure.

## BODY

### 1. Establishing Data Collection Procedures

Our pilot phase came to an end in the Spring of 1996. On April 1 1996, we completed our final round of reliability testing of all project forms (See **Appendix F**) and ended the design testing phase for data management and linking. A high-speed double-headed scanner was purchased to assist in processing approximately 2,500 mammographic encounters per week. Patient, provider and facility identifiers are double-entered by hand and linked using bar code technology and scanning. We are using a probability-based matching program to accurately assign data to patient files and for up-sequencing of multiple visits to one data record to track mammographic occurrences by breast, woman, facility, and by radiologist interpretation. We have designed all the training materials for mammography facilities and the quality assurance systems for data checking. Four field coordinators (2 permanent and 2 temporary) were hired and trained, and all mammography facilities have received several implementation and support visits by one of more of these coordinators from May 1st, 1996 through January, 1997.

To date 39 of New Hampshire's 42 mammography facilities have provided data to the NHMN. Five facilities are using our computer system for mammography data collection (See **Appendix G**), and we take data downloads from

them on a quarterly basis. Women participants continue to sign and complete the General Information Form (**Appendix F**), which is scanned at the Project office. One of the remaining facilities has suspended data collection activities due to staffing issues. We anticipate bringing it back on after these issues have been resolved. The remaining two facilities refuse to participate. **Table 1** (next page) illustrates implementation start dates and status of sites not currently contributing data to the Network.

**Table 1 New Hampshire Mammography Network Status March 1, 2000**

| Facility | Implementation Date | Type of Data Collection System |
|---|---|---|
| A | 5/28/96 | Paper |
| B | 6/10/96 | Paper |
| C | 7/1/96 | Paper |
| D | 7/1/96 | Paper |
| E | 7/8/96 | Paper |
| F | 9/3/96 | Paper |
| G | 2/2/97 | Paper |
| H | 9/23/96 | Paper |
| I | 8/1/96 | Paper |
| J | 11/1/96 | Paper |
| K | 6/3/96 | Paper |
| L | 6/3/96 | Paper |
| M | 7/2/96 | Paper |
| N | 6/ 24/96 | Computer |
| O | 9/16/97 | Computer |
| P | 9/23/96 | Computer |
| Q | 9/23/96 | Computer |
| R | 9/23/96 | Computer |
| S | 7/15/96 | Paper |
| T | 9/3/96 | Paper |
| U | 8/5/97 | Paper |
| V | 5/1/96 | Paper |
| W (2 sites) | 5/1/96 | Paper |
| X | 5/1/96 | Paper |
| Y | 11/1/96 | Paper |
| Z (2 sites) | 10/8/97 | Paper |
| AA | 10/8/97 | Paper |
| BB | 10/15/96 | Paper |
| CC | 8/5/96 | Paper |
| DD | 8/7/96 | Paper |
| EE (3 sites) | 9/3/96 | Paper |
| FF | 9/3/96 | Paper |
| GG | 9/3/96 | Paper |
| HH | 1/297 | Paper |
| II | 7/2/96 | Paper |
| JJ | 9/23/96 | Paper |
| KK | Hold | Due to staffing shortage |
| LL | REFUSED | |
| MM | REFUSED | |

An additional goal for implementation was to monitor the status of mammography facilities in their contribution of data to the project. Each facility receives a status report at quarterly intervals indicating the total volume of mammograms done, the number of women refusing to take part, the number of women not approached due to scheduling or other problems, and the amount of essential information that has not been received from that site. For each of these site-specific variables, we provide comparisons to the state aggregate, which assists the sites in determining their participation status relative to other facilities as a whole. These status reports are critical in assisting the facilities to follow-up on missing data and in identifying problem areas in the process of data collection for correction. **Appendix H** contains a sample status report used for this purpose. Upon receipt of the status reports, facilities are entered into our system for follow-up of missing data (Called our "Chase and Trace" System). Forms that are missing essential information are photocopied onto bright pink paper and are returned to the facility for completion or correction. The implementation of this system has resulted in improved completion rates on data forms at the first point of submission, and the status reports have been enormously helpful in improving the overall completeness of data contained in the NHMN.

Because the accuracy of data is so critical to the research conducted using NHMN data, we have incorporated several quality assurance measures into the process of data collection. First, the scanning technology we are using to process project forms has set parameters for acceptance or rejection of data. For example, if a woman indicates she has no breast concerns on the Patient Intake Form but goes on to describe a breast lump to the mammography technologist, the form will reviewed and verified. Staff operating the verification station for the scanner has been trained on all parameters for verification. Second, the patient registration system (where patient identifiers are double entered) automatically selects cases (10% of cases are selected at random, based on volume of mammographic encounter for each facility) for radiologist report quality assurance. For the selected cases, consent forms are copied and facilities pull the radiologist reports. The field coordinators review the text reports and complete a corresponding radiologist form. These forms are then compared with the reports submitted by the participating radiologists, and discrepancies are reviewed by our radiologist liaison. To date, there is a 96% agreement between the field coordinators' interpretation of the text reports and their completed radiologist reports, indicating that radiologists are completing their forms correctly. Our radiology liaison follows up with any radiologist using an incorrect format in completing data forms.

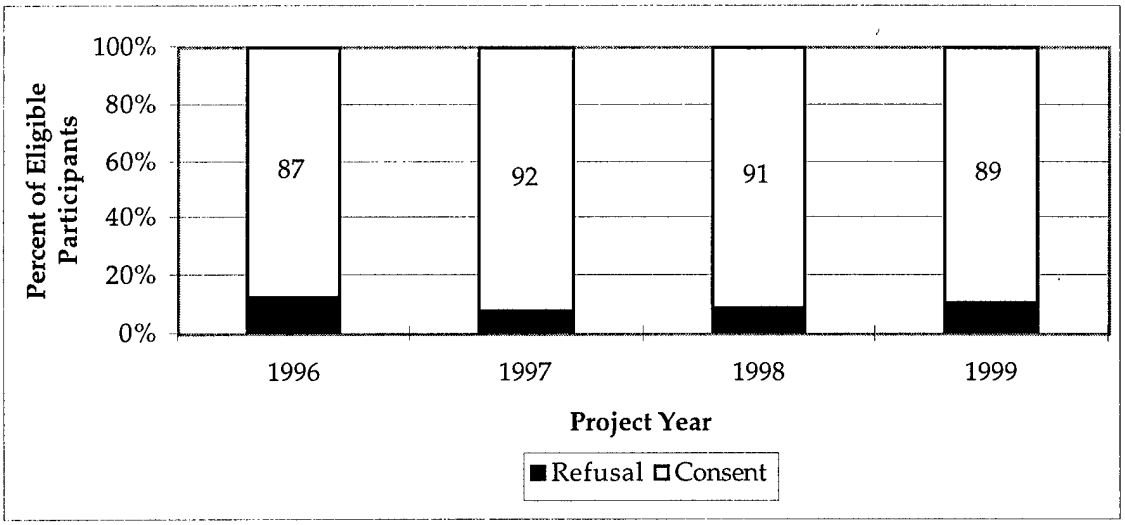
In our original proposal, we planned to contract with tumor registrars to abstract breast pathology reports at New Hampshire labs. In part, because of the funding we received for the NH Quality Assurance Project, the labs are sending their pathology reports to our Project office and they are abstracted on-site. Our pathology interpretation form is included in **Appendix I**. Quality assurance is performed by our pathology liaison (a pathologist at Dartmouth-Hitchcock Medical Center) on 25% of the abstracted pathology reports, with greater than 94% agreement

between the pathology liaison and the abstractor. Our institutional review board has given us permission to hold identifiers from breast tissue reports for six months, to allow for adequate matching with the NHMN. When this six-month period passes, identifiers are dropped from the database and anonymous data remains. We have developed and tested our matching protocols with the NH State Tumor Registry and are able to perform the linkages between women in the NHMN and the breast pathology database, which we do semi-annually.

The creation of the NHMN database, data management processes (for paper and computerized systems), and data linking for analyses have all been accomplished. These are fully described in a published paper and accompanying commentary, which are included in **Appendix I.**

**Figure 2** (below) outlines the consent and refusal rates for eligible participants over the four active years of the Project. Over 250,000 mammographic encounters have been entered into the database. The majority of women in the database are over age 50 (55%) and 45% are under age 50. Consent rates have fluctuated on a monthly basis between 87%-96% with a mean of 91%. The follow-up of missing data ranges from 1.3% - 3% (see **Figure 3,** next page). This missing data is updated in the database when the follow-up reports are returned. All sites but 1 (97.5%) are participating in our follow-up system for missing data. Those not approached (due to site-specific circumstances and those who are disabled) range from 1-3%.

**Figure 2 Volume and Status of the NHMN Database March 1, 1998-March 1, 1999 (current number of facilities = 40).**



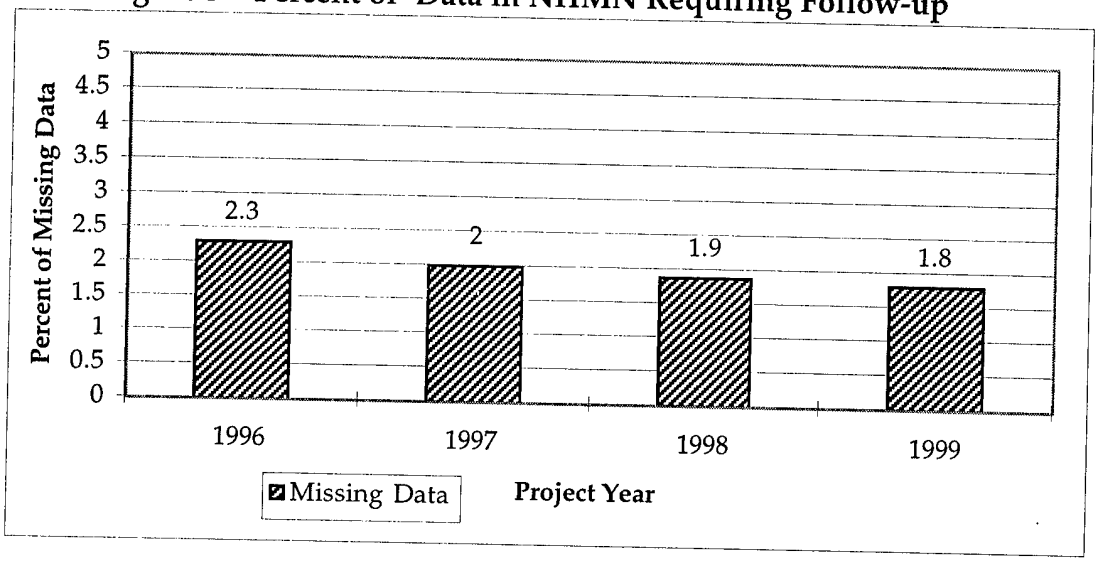

Once the creation of the database, data management processes (for paper and computer systems), and data linking for analyses were accomplished, our next challenges included generating data reports for facilities and radiologists and determining if the reporting process improved the diagnostic acumen of

participating radiologists, which brings us to the next set of tasks outlined in the original proposal.

## 2. Data Analysis and Feedback Reporting Procedures

In preparation for evaluating the impact of our reporting procedures, we defined our indices for accuracy. The following definitions have been agreed upon by the New Hampshire Mammography Network Steering Committee for purposes of conducting data analyses on accuracy.

**Figure 3 Percent of Data in NHMN Requiring Follow-up**




a. Indication for Exam

Screening Indication- A standard two view (craniocaudal and mediolateral) mammogram whose occurrence is not influenced by concerns about the presence of symptoms, positive clinical breast exam, or prior mammogram nine months ago or less (< 270 days). Indication for exam is not influenced by use of additional views, or ultrasound or any ACR assessment. Hierarchical classification for coding is: 1) Patient Intake Form (Tech.) - presence of breast concerns is no; 2) Radiologist Form - is Bilateral Asymptomatic (screening mammogram); 3) General Health Questionnaire - is a routine screening exam.

Diagnostic Indication - Mammogram (that may include non-standard additional mammographic views or supplemental ultrasound) whose occurrence is associated with any of the following: concerns about the presence of symptoms, a positive clinical breast exam or prior mammogram within 9 months (< 270 days). Hierarchical classification for coding is: 1) Patient Intake Form (Tech.) - presence of breast concerns is yes (does not include pain); 2) Radiologist Form is Diagnostic mammogram (for clinical indication), Follow-up at short interval (to evaluate stability), or Additional Views to supplement screening exam, or; 3) General Health Questionnaire - is anything BUT routine screening exam.

b. Type of Exam Performed

1a) Standard Screening Mammogram - A standard two view (craniocaudal and mediolateral) mammogram obtained for a screening indication [see above] which does not include supplementary imaging (i.e., additional mammographic views or breast ultrasound). Repeat views for sub-optimal technical quality do not change status of exam as screening.

1b) Screen Plus Mammogram - A mammogram: (1) obtained for a screening indication which includes supplementary imaging (i.e., additional mammographic views or breast US) or (2) requested to supplement a screening mammogram that occurred within 45 days.

1b) Special case - A mammogram obtained for screening indication and assessed as ACR 0 OR recommended for immediate additional evaluation (additional mammographic views or breast ultrasound), which is succeeded by a supplementary imaging encounter within 45 days should also be considered a screen plus.

Supplementary exams obtained >45 days from index screen are considered diagnostic and the preceding index mammogram is considered a standard screening mammogram.

Diagnostic Mammogram - A mammogram (which may include additional non-standard mammographic views or breast US) obtained for a diagnostic indication [see above].

Recall Rate - The percent of mammograms obtained for a screening indication that necessitate additional imaging, such as non-standard mammographic views and/or breast US.

Positive Screening Mammogram Interpretation - A screening interpretation will be considered positive: 1) if the American College of Radiology (ACR) Lexicon Code is 0 (assessment incomplete), 4 (suspicious abnormality), or 5 (highly suggestive of malignancy) OR 2) any screening mammogram interpretation (ACR Lexicon Code of 0-5) that is accompanied by recommended follow-up for any additional work-up. In practice settings wherein only completed assessments are reported (i.e., ACR 0 assessments are resolved prior to reporting, the screening mammogram will be interpreted as ACR code = 0 if there is any additional work-up performed beyond the screening mammogram).

Negative Screening Mammogram Interpretation - A screening interpretation will be considered negative if the ACR code is 1 (negative) or 2 (benign finding, negative) AND the recommended follow-up for routine mammogram is greater than 9 months (> 270 days).

Positive/Negative Screening Mammogram Interpretation - A screening interpretation will be considered positive in the first analysis, and then negative in a

repeated analysis if the ACR code is 3 (probably benign finding) AND the recommended follow-up is for less than nine months (< 270 days).

Cancer Diagnosis - An outcome is defined as cancer (or positive) if there is a histologic proved diagnosis of DCIS or invasive cancer, or NH State Tumor registry documentation for invasive cancer or DCIS within the follow-up period.

Non-Cancer Diagnosis - An outcome is defined as non-cancer (or negative) if there is a proven benign diagnosis or no pathology at the end of the follow-up period (one or two years).

Follow-up Time -12 months - The twelve-month analysis will be based on a time period of greater than 12 months from the date of the index mammogram (>365 days). The index mammogram is the first mammography encounter of the time interval under analysis.

Follow-up Time - Two Years - The two-year analysis will be based on a time period of 24 months time period from the date of the index mammogram. For the two-year analysis, two years would be substituted for one year in the analyses below (Item 10).

c. Accuracy Indicators

Positive Screen Mammogram Interpretation and, True Positive (TP), and False Positive (FP) - A positive screening mammogram interpretation is a true positive if there is a cancer diagnosis (date of diagnosis will be used for time period indicator) before the end of the follow-up period. This is regardless of the mode of detection (screening or diagnostic). A positive screening mammogram interpretation is a false positive if there is no cancer diagnosis (date of diagnosis will be used for time period indicator) before the end of the follow-up period.

Negative Screen Mammogram, True Negative (TN), and False Negative (FN) - A negative screening mammogram interpretation is a true negative if there is no cancer diagnosis before the end of the follow-up period. A negative screening mammogram interpretation is false negative if there is a cancer diagnosis date before the end of the follow-up period.

d. Analyses

Screening Interpretation Only - The initial analysis will be for screened mammograms only. In order to include all women in the analysis, women having had additional evaluations at the time of the index mammogram will be included. The mammogram interpretation for these women would be considered as ACR "0" for this analysis.

Screening Plus Additional Evaluation Interpretation (Screen-Plus) - The second analysis will be for screening mammography plus further diagnostic work-up (where a supplementary breast imaging encounter occurs within 45 days). (Infrequently more than one supplementary encounter will occur and should be linked together providing the interval between successive encounters does not exceed 45 days). For this analysis, we would use the ACR codes assigned at the end of the complete work-up process, including all radiologic studies up to, but not including, biopsy for all women.

Table 2 Illustrates the indices for calculating accuracy.

**Table 2 Indices for Calculating Accuracy**

| Mammography Result | Cancer Status* Positive | Negative | |
|---|---|---|---|
| Mammo + | TP | FP | Total Test + |
| Mammo - | FN | TN | Total test - |
| Total | Women with cancer | Women without cancer | |

Sensitivity = TP/TP + FN
Specificity = TN/FP + TN
Positive Predictive Value = TP/TP +FP
Negative Predictive Value = TN/FN +TN

* A histologically or registry proved ductal carcinoma in situ or invasive primary cancer of the breast. Lobular carcinoma in situ will be included in one analysis, then removed for a second analysis.

We have developed report formats, which were approved by the NHMN Steering Committee (sample reports using fictional data are included in **Appendix K**). The Steering Committee is composed of members of the research team, community radiologists, community pathologists, and mammography technologists. Any report that contains patient-level information is treated as confidentially as any medical record (as noted in the Confidentiality Manual). Dummy codes are generated each time a report is created to protect the identity of a receiving facility or radiologist. These codes never link radiologist participants to actual study identifiers. Confidentiality and privacy issues are outlined in a Report Handling Policy that is included in **Appendix A**.

Our initial analysis, of the performance of mammography in NH, using the above definitions, which was conducted in June of 1999 (for time period 1/1/97 to 12/31/97) to allow enough time for follow-up of breast pathology cases, indicates that prior to receiving feedback reports the statewide Sensitivity was 70.5% (95% CI-

64.7-76.3%); Specificity was 97.3% (95% CI-97.2-97.4%); Positive Predictive Value was 10.0% (95% CI-8.5-11.4%); and Negative Predictive Value was 99.9% (95% CI-99.8-99.9%). These analyses were based on 55,965 women who were receiving screening mammograms. Subsequent analyses done in September 2000 for the time period 1/1/98-12/31/98 indicate that since receiving feedback reports, these performance measures have improved. Statewide Sensitivity in this most recent analysis was 73.4% (95% CI-68.7-78.6%); Specificity was 94.9% (95% CI-94.7-95.1%); Positive Predictive Value was 6.3% (95% CI-5.5-7.1%); and Negative Predictive Value was 99.9% (95% CI-99.8-99.9%).

The work we conducted in preparing data and generating useful information for radiologists and facilities has prepared us to use this important resource for research, which will be addressed in the final section of the Body of this report.

**3. Expanded Use of the Infrastructure**

Our tasks related to this objective were to present results regionally and nationally and to coordinate with other investigators for new research studies using the data we collect. This section will specifically focus on studies conducted, underway, and those planned for which we have received peer-reviewed funding. The **Reportable Outcomes Section** (page 27) lists the presentations we have conducted regionally and nationally.

a. Completed Studies

*a.1. The 1997 New Hampshire Breast Pathology Quality Assurance Study* was funded by the State of New Hampshire Department of Health and Health Services through a cooperative agreement with the Centers for Disease (grant # U57-CCU108362-02)). Its purposes were two –fold: 1) to evaluate the diagnostic agreement among NH pathologists interpreting breast tissue and 2) evaluate the accuracy and reproducibility of DCIS grading systems. The first Project was significant in that no previous studies assessed agreement among community pathologists (not recognized experts in the field) using a representative sample of cases that reflect everyday practice.

In this study, each pathologist evaluated slides from 30 cases randomly selected from a statewide breast pathology database. The diagnostic categories used in the evaluation were benign, benign with atypia, non-invasive malignant, and invasive malignant. Twenty-six (59%) of the 44 eligible pathologists in the State participated in the slide review. Diagnostic agreement was assessed using the kappa coefficient. We found that agreement was high among pathologists for determining diagnostic category (kappa = 0.71), and was nearly perfect for benign versus malignant categories (kappa = 0.95). There was less agreement for the categories of non-invasive malignant and benign with atypia (kappa = 0.59 and 0.22, respectively). There was no apparent relationship between levels of diagnostic agreement and tissue source or perceived slide quality. In conclusion, we found that diagnostic

agreement for breast tissue specimens is high overall among community-based pathologists, but clinically relevant disagreements may occur in the assessment of non-invasive malignant diagnoses. Establishing reread policies for certain diagnostic categories may reduce the possibility that diagnostic misclassification will lead to over- or under treatment. The high diagnostic reproducibility for malignant lesions of breast suggests that it is unnecessary for a central review of these lesions in national cancer trials. Two publications that describe this project are included in **Appendix C.**

*a.2. The 1998 New Hampshire Breast Pathology Quality Assurance Study* was also funded by the State of New Hampshire Department of Health and Health Services through a cooperative agreement with the Centers for Disease (grant # U57-CCU108362-02). The second breast pathology Quality Assurance Project was significant because many pathologists have attempted to describe the different types and patterns of non-invasive carcinomas of ductal origin (DCIS) (5). The poorly defined criteria for differentiation of these patterns have mainly concentrated on the architectural features and the presence or absence of necrosis (6). Unless the diagnostic reproducibility of these different DCIS grades among every day, practicing pathologists can be determined, the usefulness of such a grading system nationwide will remain unknown and its impact in treatment decisions limited. In this study, seven non-expert pathologists in New Hampshire and three experts evaluated forty slides of DCIS according to three internationally recognized classifications. Twenty slides were re-interpreted by each non-expert pathologist. Diagnostic accuracy (non-experts as compared to experts) and reproducibility were evaluated using inter- and intra-rater techniques (kappa statistic).

We found that final DCIS grade and nuclear grade were most accurately reported among non-expert pathologists using HL (Holland) (7) (kappas = 0.53 and 0.49 respectively), as compared to LA (LaGios) (8) and VN (Van Nuyes) (9) (kappas =0.29 and 0.35 respectively for both classifications). An intermediate DCIS grade was most accurately assessed using HL and LA, and a high grade (Group 3) using VN. Diagnostic reproducibility was highest using HL (kappa=0.49). The VN interpretation of necrosis (present or absent) was more accurately reported than the LA criteria (extensive, focal or absent) (kappas = 0.59 and 0.45 respectively) but reproducibility of each was comparable (kappas=0.48 and 0.46 respectively). Intra-rater agreement was high overall. In conclusion, when we compared all three classifications, final DCIS grade was reported best using HL. Nuclear grade (cytodifferentiation) using HL and the presence or absence of necrosis was the criteria most accurately and reproducibly diagnosed. Establishing one internationally approved set of interpretive definitions, with acceptable accuracy and reproducibility among both pathologists with and without expertise in breast pathology interpretation, will assist researchers in evaluating treatment effectiveness and characterizing the natural history of DCIS breast lesions. A publication that describes this project in detail is included in **Appendix C.**

*a.3. The Characteristics of the Screened Population in New Hampshire* - New Hampshire (NH) is one of two states that have developed a population-based mammography registry. After collecting data for 20 months, we have characterized the women who are receiving mammography in NH and the imaging that is done. The database contained almost 110,000 mammographic encounters representing 101,679 NH women when this analysis was done, who range in age from 18 to 97 with a mean of 56.7 years (SD=10.91). Education levels are high with 92% having a high school education and 59% with some college. Forty-six percent report their primary insurance is private, 29% report HMO/PPO coverage, and 25% receive federal health care assistance. Risk factors represented in the database include (categories not mutually exclusive) advancing age (60% over age 50), hormone replacement therapy use by menopausal women (40.6%), and a family history of breast cancer (29%). Penetration of mammography relative to the NH population is higher for younger age groups (40-48% for those aged 44-64) than older age groups (34-39% for those aged 65-84). The majority of mammographic encounters are routine screening exams (86%), often interpreted as negative or normal with benign findings (88%). Use of comparison films to interpret either diagnostic or screening mammography occurred in 86% of encounters. We have matched 3,877 breast pathology records to these mammographic encounters. The distribution of pathology outcomes for diagnostic exams was very similar to that for screening exams (approximately 65% benign, 17% invasive breast cancer, and 6% non-invasive breast cancer). Overall, we have designed a system that is well accepted by the NH community. Challenges include careful monitoring of data for coding errors, and a limitation of linking variables in mammography and pathology data. Data represented in this registry are a critical resource for research in mammographic screening and breast cancer early detection. The publication from this work is included in **Appendix B**.

*a.4. The Performance of Mammography in New Hampshire* – This analysis was recently completed with Dept. of Defense support through the original grant. The purpose of this analysis was to describe the practice of mammography in a statewide population. Mammography data on 47,651 screening and 6,152 diagnostic examinees, from the time period 11/1/96 to 10/31/97 were linked to 1,572 pathology results. Mammography outcomes were based on BI-RADS (10) assessments and recommendations reported by the interpreting radiologist. The consistency of BI-RADS recommendations was also evaluated. Our results indicated that screening mammography had a sensitivity of 72.4% (95%CI: 66.4 - 78.4%), specificity of 97.3% (95%CI: 97.2 -97.4%), and positive predictive value (PPV) of 10.6% (95%CI: 9.0 - 12.2%). Diagnostic mammography had higher sensitivity - 78.1% (95%CI: 71.9 - 84.3%), lower specificity - 89.3% (95%CI: 88.5 - 90.1%), and better PPV - 17.1% (95%CI; 14.4 -19.8%). The cancer detection rate of screening was 3.3 per 1000 with a biopsy yield of 22%, while the interval cancer rate was 1.2 per 1000. Nearly 80% of screen detected invasive malignancies were node negative. The recall rate for screening was 8.3%. Ultrasonography was utilized in 3.5% of screening, and 17.5% of diagnostic encounters. BI-RADS recommendations were generally consistent, except for probably benign assessments. In conclusion, we found that the sensitivity

of screening mammography in our population- based sample is lower than expected, although other performance indicators are commendable. BI-RADS probably benign assessments are commonly misused. This work has been accepted for publication in Radiology. The In Press version of the manuscript is included in **Appendix L.**

*a.5. Performance of Screening Mammography Among Women with and without a First-degree Relative with Breast Cancer* – This analysis was recently completed with Dept. of Defense support through the original grant and was a collaborative multi-center effort. We conducted this analysis to determine the performance of screening mammography in women with a first-degree family history of breast cancer compared to women without of similar age. This is important because women with a family history of breast cancer are often recommended to undergo regular screening mammography beginning at a younger age. Few studies have evaluated the performance of screening mammography among women at increased risk of breast cancer. Our study design was cross-sectional and involved seven mammography registries in San Francisco, Seattle, New Hampshire, New Mexico, Vermont, Washington state and Colorado. Participants included 389, 533 women aged 30 to 69 years referred for screening mammography from April 1985 to November 1997. Data amassed for the analysis included: breast cancer risk factors, first mammography screening examination interpretation, follow-up of abnormal and normal mammography to determine occurrence of invasive cancer or ductal carcinoma in situ by linkage to either a pathology database, the Surveillance, Epidemiology, and End Results program or to a state tumor registry.

We found that the rate of cancer per 1000 examinations increased with age and was higher among women with a family history of breast cancer (3.2 for ages 30-39 [95% CI 1.7, 4.6], 4.7 for ages 40-49 [95% CI 3.8, 5.7], 6.6 for ages 50-59 [95% CI 5.3, 8.0], 9.3 for ages 60-69 [95% CI 7.5, 11.1]; Chi-square for trend P= .001) compared with those without (1.6 for ages 30-39 [95% CI 1.2, 2.0], 2.7 for ages 40-49 [95% CI 2.4, 2.9], 4.6 for ages 50-59 [95% CI 4.1, 5.0], 6.9 for ages 60-69 [95% CI 6.3, 7.5]; Chi-square for trend P= .001). The sensitivity of mammography increased with age among women with a family history of breast cancer (63.2% for ages 30-39 [95% CI 41.5, 84.8], 70.2% for ages 40-49 [95% CI 61.0, 79.5], 81.3% for ages 50-59 [95% CI 73.3, 89.3], 83.8% for ages 60-69 [95% CI 76.8, 90.9]; Chi-square for trend P= .001) and those without (69.5% for ages 30-39 [95% CI 57.7, 81.2], 77.5% for ages 40-49 [95% CI 73.3, 81.8], 80.2% for ages 50-59 [95% CI 76.5, 83.9], 87.7% for ages 60-69 [95% CI 84.8, 90.7]; Chi-square for trend P= .001) but was similar for each decade of age irrespective of family history status. In conclusion, having a first-degree relative with a history of breast cancer was associated with cancer detection rates similar to women a decade older without a family history. The sensitivity of screening mammography was primarily influenced by age. The In Press version of the manuscript is included in **Appendix M.**

b. Studies Currently Underway

*b.1. Assessing and Improving Interval Mammography Adherence* - In 1998, the American Cancer Society funded a study to evaluate the characteristics of women aged 50 and older who do and do not adhere to interval mammography screening. We used data from the NHMN to identify those adhering and not adhering to interval screening. The rationale for this study is that annual mammography screening in women aged 50 and older is associated with a 30% reduction in breast cancer mortality. Self-reported interval screening adherence rates range from 21-84%. Many factors appear to affect adherence to screening, such as knowledge, beliefs, risk perception, and anxiety or worry. Much less is known about how these factors interact to promote or discourage screening behavior. We used the NH Mammography Network (NHMN) to evaluate the extent to which NH women in this age group adhered to interval screening and found an interval adherence rate of 24% in women aged 50 and older.

The specific aims of this investigation include to:

1. Assess knowledge and beliefs about breast cancer, mammography screening, objective and subjective risk, and anxiety in women (aged 50+) who do and do not adhere to interval screening;
2. Assess the impact of an education and counseling program on adherence rates;
3. Conduct subgroup analyses on women with and without a family history of breast cancer;
4. Conduct subgroup analyses on women who do and do not experience a false positive screening mammogram.

**Appendix N** includes the survey we are using to evaluate main study measures, as described above. The study will be conducted in three phases. Phase I is now complete and involved an observational study where we identified and recruited adherers (n=320) and non-adherers (n=320), defined as women aged 50 and older who had one screening mammogram and no other exam within 24 months. We collected baseline data on these women and have evaluated their demographic characteristics, which are outlined in **Tables 3 and 4** (next two pages).

Phase II is a randomized controlled trial of the impact of a telephone counseling intervention. In this phase, the non-adherers were randomized to receive telephone counseling (n=160) (based on motivational interviewing/ Transtheoretical Model of Prochaska/DiClemente) (12) or to a comparison group (n=160). The intervention is customized to the women based on the stage of change they are in when considering mammography screening. The stages include contemplation, determination, action, maintenance, and relapse. Women receive individualized information to assist in overcoming barriers, counselors help women identify "triggers" that might cause them to miss a mammography appointment and assist them in learning from their experiences. They additionally

provide empathy and support for additional attempts to obtain a screening mammogram.

**Table 3 Demographic Characteristics of NH Women Aged 50 and Older Who Do and Do Not Adhere to Interval Mammography Screening**

| Characteristics | Adherers (n=295) | | Non-Adherers (n=295) | | |
|---|---|---|---|---|---|
| | N | % | N | % | p-value |
| Age | | | | | 0.95 |
| <50 | 23 | 7.8 | 37 | 12.5 | |
| 50-59 | 167 | 56.6 | 159 | 53.9 | |
| 60-69 | 95 | 32.2 | 81 | 27.5 | |
| 70+ | 10 | 3.4 | 18 | 6.1 | |
| Education | | | | | 0.71 |
| Less than high school | 8 | 2.8 | 15 | 5.3 | |
| High school grad | 94 | 32.5 | 93 | 32.8 | |
| Associate degree | 101 | 35.0 | 96 | 33.8 | |
| College Grad | 40 | 13.8 | 41 | 14.4 | |
| Post graduate | 46 | 15.9 | 39 | 13.7 | |
| Marital status | | | | | 0.88 |
| Single | 14 | 5.0 | 16 | 5.7 | |
| Married | 222 | 78.5 | 214 | 75.6 | |
| Separated | 2 | 0.7 | 1 | 0.4 | |
| Divorced | 23 | 8.1 | 25 | 8.8 | |
| Widowed | 22 | 7.8 | 27 | 9.5 | |
| Type of Health Insurance | | | | | |
| Any | | | | | 0.001 |
| No | 1 | 0.3 | 14 | 4.8 | |
| Yes | 293 | 99.7 | 276 | 95.2 | |
| Private | | | | | 0.72 |
| No | 126 | 42.9 | 120 | 41.4 | |
| Yes | 168 | 57.1 | 170 | 58.6 | |
| Medicare | | | | | 0.13 |
| No | 253 | 86.1 | 236 | 81.4 | |
| Yes | 41 | 13.9 | 54 | 18.6 | |
| Medicaid | | | | | 0.42 |
| No | 290 | 98.6 | 288 | 99.3 | |
| Yes | 4 | 1.4 | 2 | 0.7 | |
| HMO/PPO | | | | | 0.05 |
| No | 206 | 70.1 | 224 | 77.2 | |
| Yes | 88 | 29.9 | 66 | 22.8 | |

**Table 4 Additional Characteristics of NH Women Aged 50 and Older Who Do and Do Not Adhere to Interval Mammography Screening**

| Characteristics | Adherers (n=295) | | Non-Adherers (n=295) | | |
|---|---|---|---|---|---|
| | N | % | N | % | p-value |
| Distance from Facility Used | | | | | 0.61 |
| 0-10 miles | 196 | 66.4 | 196 | 66.4 | |
| 11-20 miles | 73 | 24.8 | 73 | 24.8 | |
| 21-30 miles | 16 | 5.4 | 15 | 5.1 | |
| 31-40 miles | 2 | 0.7 | 6 | 2.0 | |
| 41+ miles | 8 | 2.7 | 5 | 1.7 | |
| Family History of Breast Cancer | | | | | 0.17 |
| 1st Degree Relative | 44 | 14.9 | 29 | 9.8 | |
| Other Relatives | 37 | 12.5 | 39 | 13.2 | |
| No Family History of breast cancer | 214 | 72.5 | 227 | 77.0 | |
| Parity | | | | | 0.03 |
| 0 | 16 | 5.9 | 11 | 4.2 | |
| 1 | 15 | 5.5 | 26 | 9.9 | |
| 2 | 115 | 42.1 | 83 | 31.4 | |
| 3 | 70 | 25.6 | 70 | 26.5 | |
| 4+ | 57 | 20.9 | 74 | 28.0 | |
| Age at Menarche | | | | | 0.95 |
| < 11 years | 16 | 5.6 | 16 | 5.6 | |
| 11 years | 51 | 17.7 | 50 | 17.6 | |
| 12 years | 71 | 24.7 | 80 | 28.2 | |
| 13 years | 90 | 31.3 | 80 | 28.2 | |
| 14 years | 32 | 11.1 | 31 | 10.9 | |
| 15+ years | 28 | 9.7 | 27 | 9.5 | |
| Type of Menopause | | | | | 0.78 |
| Natural | 125 | 58.7 | 120 | 55.8 | |
| Surgical | 86 | 40.4 | 92 | 42.8 | |
| Radiologically Induced/Other | 2 | 0.9 | 3 | 1.4 | |
| HRT Use | | | | | 0.02 |
| No | 107 | 48.6 | 134 | 59.8 | |
| Yes | 113 | 51.4 | 90 | 40.2 | |
| Breast Density Assessment | | | | | 0.65 |
| Fat | 62 | 21.2 | 58 | 19.9 | |
| Scattered | 142 | 48.5 | 154 | 52.7 | |
| Heterogeneously Dense | 73 | 24.9 | 69 | 23.6 | |
| Extremely Dense | 16 | 5.5 | 11 | 3.8 | |

In Phase III, we will be collecting post-intervention data on both groups at 6 (completed), 18 (completed) and 30 months. The intervention was developed, pilot tested and implemented in fall of 99 (1st counseling call). The second intervention

call is scheduled for this fall (2000) and the final set of post intervention measures will be collected after the first of the year (2001). We are currently linking survey data to NHMN demographic, objective risk, and mammographic history data to begin to evaluate the effect of the intervention on women's screening behaviors. This study will be completed in July of 2001 and will help us: 1) understand how much and what type of anxiety promotes or inhibits screening behavior; and 2) develop a profile of women at risk of non-adherence to screening so that we can disseminate this information to primary care providers.

**c. Future Work**

c. Recently Funded Studies

*c.1. Hormone Replacement Therapy and Breast Cancer* - We received funding in April 2000 to evaluate the impact of hormone replacement therapy on mammography performance, breast cancer incidence, tumor prognostic characteristics and health-related quality of life by following a well-defined population-based cohort of women who use mammography. Although the results of case-control and follow-up studies (13, 14) suggest that hormone replacement therapy modestly increases breast cancer risk, most studies have been unable to account adequately for frequency of mammographic screening. This is an important limitation because more frequent use of mammography screening among women who maintain hormone replacement prescriptions through regular physician visits may lead to increased detection of breast cancer relative to women who do not use hormone replacement therapy. Our proposed study, which is based on the New Hampshire Mammography Network (NHMN), overcomes this limitation. The NHMN comprises over *152,000* women who have completed a baseline survey including data on breast cancer risk factors and use of hormone replacement therapy. These women have also provided permission to link medical, radiologic and pathology data, and have consented to further contact for research purposes. Through NHMN we have already identified 74,200 women who are peri- or post menopausal including approximately 26,700 current HRT users. We will follow these women prospectively to ascertain new cases of breast cancer. Our primary specific aims and related hypotheses are as follows:

1. To evaluate the impact of hormone replacement therapy on the sensitivity and specificity of screening mammography and on the proportion of uninterpretable mammograms and consequent use of other imaging procedures (e.g., breast ultrasound).

   H1: Current hormone replacement therapy use will be associated with decreased mammographic sensitivity and specificity, increased frequency of uninterpretable mammograms, and increased use of other imaging procedures.

2. To evaluate the relationship between hormone replacement therapy (especially combination therapies) and breast cancer incidence.

H2: Current use of hormone replacement therapy, long term use, and recent past use will be associated with increased risk of invasive breast cancer incidence after frequency of mammography is taken into account.

3. To compare breast cancer detection characteristics (e.g., proportion screen-detected versus interval cancers) and breast tumor prognostic characteristics (e.g., TNM stage, tumor grade, axillary lymph node status, and estrogen receptor status) according to hormone replacement therapy use.

H3: Hormone replacement therapy will be associated with increased rates of interval versus screen detected cancers and there will be differences in tumor characteristics (TNM stage, tumor grade, axillary lymph node status, and estrogen receptor status) according to history of hormone replacement therapy use.

As more women consider taking HRT and other post menopausal therapies (e.g., raloxifene) to prevent osteoporosis and other diseases, it is imperative to document the impact of HRT on health-related quality of life in a broad spectrum of women. Therefore a secondary aim and hypothesis is as follows:

4. To compare health-related quality of life in hormone replacement therapy users and non-users and among women with and without breast cancer.

H4: Hormone replacement therapy will be associated with increased health-related quality of life.

Results from the proposed study will be of particular relevance to radiologists who interpret mammograms, and to women and their health care providers, who must balance the complex issues of disease risk and health-related quality of life when deciding whether or not to use hormone replacement therapy. The survey we have developed and are currently pilot testing is included in **Appendix O**.

*c.2. Strategic Studies on Breast Cancer Surveillance* - We received funding in April 2000 to both support and expand NHMN activities as well as to conduct five special projects. The objectives of this proposal are to continue to expand our research capacity and to conduct special studies over the next five years. The special studies will use current and expanded data resources to enhance our understanding of breast cancer detection processes.

Our first three aims address continuation and expansion of current NHMN activities in New Hampshire (NH), while the subsequent aims address specific research goals. The first two of these, Aims 4 and 5, will be based in NH, while Aims 6-8 will use pooled data from three geographically defined mammography registries (NH, Vermont-VT, and North Carolina-NC). By pooling data from the

three registries, we will increase the number of cancer cases for analyses and will include a more diverse population in our research.

- Our infrastructure enhancing aims are to:

1. Continue and refine current NHMN procedures, including data collection, pathology linkages, feedback reports to mammography facilities and radiologists, and data submissions to a centralized statistical coordinating center;

2. Implement a process to obtain follow-up information on NH women with a mammographic abnormality who did not obtain follow-up care or whose follow-up care was not ascertained by the NHMN;

3. Implement a process to obtain information characterizing the ultimate pathway of breast cancer discovery in NH women with false negative mammograms, including what motivated them to seek follow-up care.

- Our special studies aims and related hypotheses are to:

4. Compare actual risk, perceived risk, and anxiety traits in a population-based sample of unscreened NH women to screened women in the NHMN database.

H4-a: Risk factors in screened and unscreened women will be similar within age strata.

H4-b: Compared to screened women, unscreened women will be more likely to have: 1) high anxiety character traits/high risk perceptions and 2) low anxiety character traits/low risk perceptions.

5. Evaluate the influence of menstrual cycle phase on breast density and mammographic performance.

H5-a: Increased breast density will be associated with the luteal phase of the menstrual cycle.

H5-b: Mammography accuracy will be lower in the luteal phase of the menstrual cycle.

6. Determine whether benign breast biopsy characteristics (biopsy type, number of breast biopsies and biopsy outcome) are related to mammographic accuracy.

H6-a: Mammographic accuracy is inversely associated with the number of previous breast biopsies;

H6-b: Mammographic accuracy is lower with a history of more invasive breast biopsies.

H6-c: Mammographic accuracy is highest in women with a previous history of lobular carcinoma in situ, ductal carcinoma in-situ, and atypical ductal hyperplasia.

7. Develop breast cancer risk prediction models for invasive and non-invasive (in situ) breast cancer using screened women. These models will include breast density as a marker for breast cancer risk and breast cancer screening performance;

8. Develop a longitudinal model of mammography/health states defined by screening compliance, mammography outcomes, follow-up and disease outcomes in individual subjects, and determine predictors for transitions between these states. This aim will expand on Aim 7 and will include an evaluation of mammography-related predictors for all cause mortality and breast cancer mortality.

*c.3. Understanding Variability in Community Mammography* – On September 1st, 2000, we received funding from the Agency for Research in Health Quality (ARHQ). The goal of this research is to identify reasons for variability in the interpretation of mammograms. Though previous studies have shown marked interpretive variation, they did not explain why it occurs and they used test sets that do not necessarily reflect what occurs in day-to-day community practice. This community-based multi-center observational study will utilize a unique collaboration among three geographically distinct breast cancer surveillance programs in Washington, New Hampshire, and Colorado. This collaboration will allow us to accumulate breast cancer outcome and interpretive performance data on more than 500,000 mammograms from 91 facilities and 279 radiologists.

We will evaluate potential factors influencing the accuracy and recall rate of mammography using a structured conceptual framework that separates characteristics of the radiologists from those of the facility and community environment. Gaining a better understanding of the how individual radiologists and their practice environment account for variation will help identify ways to improve mammography. Our overarching hypothesis is that the fiscal environment, legal environment, individual radiologist characteristics and practice environment influence variability in the accuracy of mammography and the likelihood of having patients recalled for additional evaluation.

Our specific aims are to:

1. Evaluate the influence of radiologist level characteristics on variation among radiologists' mammography recall rates and accuracy (sensitivity, specificity, or positive predictive value). Salient radiologist level variables will include:
    a. Fiscal incentives (e.g., bonus incentive package, salary structure);
    b. Legal factors (e.g., perceived or actual high levels of medical malpractice activity, past personal experience with malpractice suits)

c. Personal characteristics of the radiologist (e.g., experience interpreting mammograms, level of comfort dealing with ambiguity in medicine, concern over missing a cancer, reports on mammography interpretation

2. Evaluate the influence of facility level characteristics on variation among mammography facilities I mammography recall rates and accuracy (sensitivity, specificity, or positive predictive value). Salient facility level variables will include:
    a. Fiscal environment (e.g., for-profit status, predominantly fee-for-service payer mix)
    b. Legal environment (e.g., high density of medical malpractice lawyers in the area, high density of medical malpractice cases)
    c. Community practice environment (e.g., extent of managed care market penetration, density of mammography facilities and radiologists, availability of on-site diagnostic services.

3. To explore, using hierarchical modeling techniques, the extent to which fiscal, legal, clinical and personal characteristics of radiologists and facilities could be varied to lower the recall rates for community-based mammography while maintaining high levels of accuracy (sensitivity, specificity, or positive predictive value).

## KEY RESEARCH ACCOMPLISHMENTS

Breast Pathology Reproducibility Studies

- Diagnostic agreement for breast tissue specimens is high overall among community-based pathologists (kappa = 0.71), and was nearly perfect for benign versus malignant categories (kappa = 0.95).

- Clinically relevant disagreements may occur in the assessment of non-invasive malignant diagnoses (kappa = 0.59 and 0.22, respectively).

- Establishing reread policies for certain diagnostic categories may reduce the possibility that diagnostic misclassification will lead to over- or under treatment.

- The high diagnostic reproducibility for malignant lesions of breast suggests that it is unnecessary for a central review of these lesions in national cancer trials.

- Final DCIS grade was reported best using the Holland Classification system.

- Nuclear grade (cytodifferentiation) using Holland and the presence or absence of necrosis were the criteria most accurately and reproducibly diagnosed.

- Establishing one internationally approved set of interpretive definitions, with acceptable accuracy and reproducibility among both pathologists with and without

expertise in breast pathology interpretation, will assist researchers in evaluating treatment effectiveness and characterizing the natural history of DCIS breast lesions.

Population-based Studies on Mammography

- Screening mammography had a sensitivity of 72.4% (95%CI: 66.4 - 78.4%), specificity of 97.3% (95%CI: 97.2 -97.4%), and positive predictive value (PPV) of 10.6% (95%CI: 9.0 -12.2%), which is lower than reported elsewhere.

- Diagnostic mammography had higher sensitivity - 78.1% (95%CI: 71.9 -84.3%), lower specificity - 89.3% (95%CI: 88.5 - 90.1%)), and better PPV - 17.1% (95%CI; 14.4 - 19.8%).

- The cancer detection rate of screening was 3.3 per 1000 with a biopsy yield of 22%, while the interval cancer rate was 1.2 per 1000.

- Nearly 80% of screen detected invasive malignancies were node negative. The recall rate for screening was 8.3%.

- Ultrasonography was utilized in 3.5% of screening, and 17.5% of diagnostic encounters.

- BI-RADS probably benign assessments are commonly misused.

- Penetration of mammography relative to the NH population is higher for younger age groups (40-48% for those aged 44-64) than older age groups (34-39% for those aged 65-84).

- The majority of mammographic encounters are routine screening exams (86%), often interpreted as negative or normal with benign findings (88%).

- Use of comparison films to interpret either diagnostic or screening mammography occurred in 86% of encounters.

- The distribution of pathology outcomes for diagnostic exams was very similar to that for screening exams (approximately 65% benign, 17% invasive breast cancer, and 6% non-invasive breast cancer).

- Rate of cancer per 1000 examinations increases with age and is higher among women with a family history of breast cancer (3.2 for ages 30-39 [95% CI 1.7, 4.6], 4.7 for ages 40-49 [95% CI 3.8, 5.7], 6.6 for ages 50-59 [95% CI 5.3, 8.0], 9.3 for ages 60-69 [95% CI 7.5, 11.1]; Chi-square for trend P= .001) compared with those without (1.6 for ages 30-39 [95% CI 1.2, 2.0], 2.7 for ages 40-49 [95% CI 2.4, 2.9], 4.6 for ages 50-59 [95% CI 4.1, 5.0], 6.9 for ages 60-69 [95% CI 6.3, 7.5]; Chi-square for trend P= .001).

- The sensitivity of mammography increases with age among women with a family history of breast cancer (63.2% for ages 30-39 [95% CI 41.5, 84.8], 70.2% for ages 40-49

[95% CI 61.0, 79.5], 81.3% for ages 50-59 [95% CI 73.3, 89.3], 83.8% for ages 60-69 [95% CI 76.8, 90.9]; Chi-square for trend P= .001) and those without (69.5% for ages 30-39 [95% CI 57.7, 81.2], 77.5% for ages 40-49 [95% CI 73.3, 81.8], 80.2% for ages 50-59 [95% CI 76.5, 83.9], 87.7% for ages 60-69 [95% CI 84.8, 90.7]; Chi-square for trend P= .001) but is similar for each decade of age irrespective of family history status.

**REPORTABLE OUTCOMES**

- **Publications**

Carney PA, Poplack SP, Wells WA, Littenberg B. Development and Design of a Population-Based Mammography Registry: The New Hampshire Mammography Network American Journal of Roentgenology, 1996;167(August):367-372.

Ballard-Barbash R, Taplin SH, Yankaskas BC, Ernster VL, Rosenberg RD, Carney PA, Barlow WE, Geller BM, Kerlikowske K, Edwards BK, Lynch CF, Urban N, Chrvala CA, Key CR, Poplack SP, Worden JK, Kessler LG. Breast Cancer Surveillance Consortium: A national mammography screening and outcomes database. American Journal of Roentgenology, 1997;169:1001-1008.

Wells WA, Carney PA, Eliassen MS, Tosteson AN, Greenberg ER. A Statewide Study of Diagnostic Agreement in Breast Pathology. JNCI, 1998;90(2):142-144.

Carney, PA, Eliassen MS, Wells WA, Swartz WG. Can We Improve Breast Pathology Reporting Practices? A Community-Based Breast Pathology Quality Improvement Program In New Hampshire: Journal of Community Health, 1998;23(2):85-98.

Dole G, Carney, PA. Characteristics of Underserved Women who Did or Did Not Use a Free/Low Cost Voucher as part of a Mammography Screening Program. Journal of Cancer Education, 1998;73:102-107.

Burgess KA, Harwood B, Robinson M, Carney, PA. Mammographer Participation in a Database Research Study. Radiologic Technology, 1999;70(5):453-460.

Carney PA, Goodrich ME, O'Mahony DM, Tosteson AN, Eliassen MS, Poplack SP, Harwood BG, Burgess KA, Berube BT, Wells WA, Ball J. Mammography in New Hampshire: Characteristics of the women and the exams they receive. Journal of Community Health, 2000:25(3):183-198.

Wells WA, Carney PA, Eliassen MS, Grove MR, Tosteson AN. Agreement with Expert and Reproducibility of Three Classification Schemes for Ductal Carcinoma-in-situ. Am J. Surg Path, 2000:24(5):641-659.

Carney PA, Geller BM, Moffett H, Ganger M, Sewell M, Barlow WE, Taplin, SH, Sisk C, Ernster VL, Wilke HA, Yankaskas B, Poplack SP Urban N, West MM, Rosenberg, RD, Michael S, Mercurio TD, Ballard-Barbash R. Current Medico-legal and

Confidentiality Issues in a Large Multi-center Research Program: The National Cancer Institute's Breast Cancer Surveillance Consortium. American Journal of Epidemiology. (In Press).

Poplack SP, Tosteson AN, Grove M, Wells WA, Carney PA. The Practice of Mammography in 53,803 Women from the New Hampshire Mammography Network. Radiology (In Press).

Kerlikowske K, Carney PA, Geller B, Mandelson MT, Taplin SH, Malvin K, Ernster V, Rosenberg R, Cutter G, Ballard-Barbash R. Performance of Screening Mammography Among Women with and without a Family History of Breast Cancer. Annuls of Internal Medicine, (In Press).

- **Presentations**

Carney PA, O'Donnell J, O'Donnell L, Jackson, R. (March, 1996). Enhancing physicians' breast and cervical cancer counseling and communication skills. American College of Preventive Medicine, American Teachers of Preventive Medicine Conference. Prevention 96 Dallas, Texas.

Jackson R, Carney PA, Dube C. (October, 1996). Programs for Breast and Cervical Cancer Prevention: From Medical School to Practice. American Association for Cancer Education 50th Anniversary Meeting, Chicago, IL.

Carney PA, Wells WA, M. Eliassen MS, Tosteson A, Greenberg ER. (September, 1997). A Statewide Study of Diagnostic Agreement in Breast Pathology. 1997 CDC Cancer Conference, Integrating Public Health Programs for Cancer Control.

Carney PA. (December, 1994) The NH Mammography Registry - Interactions with Community Hospitals. Proceedings at the Dartmouth-Hitchcock Medical Center, Clinical Trials in Rural Populations. Lebanon, NH

Carney PA. (October, 1995) The New Hampshire Mammography Network 2nd Annual Breast Cancer Symposium, Norris Cotton Cancer Center, Lebanon, NH.

Carney PA. (April, 1996). The New Hampshire Mammography Network. Mammography: A Clear View: A Professional Seminar for Radiologists and Mammography Technologists. Concord, NH.

Carney PA, Wells WA, Eliassen MS, Tosteson AN, Greenberg ER. (June, 1996). A Statewide Study of Diagnostic Agreement in Breast Pathology. Era of Hope Meeting, US Dept. of Defense. Washington DC.

Carney PA. (April, 1997). Screening Controversies in Women 40-49: What are the Sources of Confusion? Spring Forward with Mammography. Concord Hospital, Concord, NH.

Harper D, Carney PA. (November, 1998). Breast and Cervical Cancer Screening in New Hampshire. OB/GYN Grand Rounds. Dartmouth-Hitchcock Medical Center, Lebanon, NH.

Carney PA. (November, 1998). Medico-legal Issues in Confidentiality and Data Security: The Experience of the NCI Breast Cancer Surveillance Consortium. Confidentiality and Biostatistics Section. American Public Health Association Annual Meeting. Washington, DC.

Carney PA. (May, 1999). The New Hampshire Breast Cancer Surveillance Project. Medical Grand Grounds, Exeter Hospital, Exeter, NH.

Carney, PA (June, 1999). Diagnostic Agreement in Breast Pathology Interpretation. Women and Cancer: The Irene Bradley Parker Conference. Norris Cotton Cancer Center, Lebanon, NH.

Carney PA. (August, 1999). Medico-legal Issues in Confidentiality and Data Security: The Experience of the NCI Breast Cancer Surveillance Consortium. American Statistical Association Meeting, Washington, DC.

Carney PA , Poplack SP, Tosteson AN, Grove M, Wells WA, Goodrich, M.. (June, 2000) The Practice of Mammography in New Hampshire. Era of Hope Meeting, US Dept. of Defense, Atlanta, GA.

Carney PA (August, 2000) Strategic Studies in Breast Cancer Surveillance: The Measure of a Mountain of Data. Norris Cotton Cancer Center Grand Rounds.

Carney PA (September, 2000) Strategic Studies in Breast Cancer Surveillance: The NH Mammography Network. American Cancer Society Regional Board of Directions Meeting. New England Division.

- **Related Research Funded**

1996-1998 A Breast Pathology Quality Improvement Project - State of N.H. Division of Public Health and the Centers for Disease Control (U57-CCU108362) ($120,000).

1998-2001 Assessing and Improving Interval Mammography Adherence - American Cancer Society (CRTG-98-280-01-CCE) ($450,000); PI-P. Carney.

1998-2000 Anxiety, Risk and Breast Cancer Screening. National Institutes of Health - Shannon Award (R55 NRO 4556-01) ($100,000); PI-P. Carney.

1999-2000 Breast Cancer Surveillance in New Hampshire. National Cancer Institute ($99,701); PI-P. Carney.

2000-2004 Hormone Replacement therapy and Breast Cancer - National Cancer Institute (R01-CA080888-01A1 - $3,696,284); PI-P. Carney.

2000-2004 Strategic Studies in Breast Cancer Detection and Surveillance - National Cancer Institute (1 U01 CA86082-01) Breast Cancer Surveillance Consortium Expansion - $3,130,434); PI-P. Carney.

2000-2003 Understanding Variability in Community Mammography - Agency for Health Care Policy and Research (R01 - $1,974,476, PI - JG Elmore PI; Co-PI - P. Carney.

## CONCLUSIONS

We have accomplished all our tasks and goals for the Project. Our greatest challenges were implementing 40 mammography facilities, insuring that complete and accurate data are collected from all participating sites, and designing a system to automatically produce reports for participating radiologists and mammography facilities. We now have enough data in the registry to develop additional manuscripts, adding to the eleven that have already been or are about to be published; two additional manuscripts have been drafted and approved by our steering committee. The first is a comparison of risk factors in women with screen versus interval detected breast cancers. The second reports on follow-up recommendations and outcomes of mammography in the NHMN. These manuscripts are currently in development. We have additionally presented at more than 15 regional or national meetings on various topics related to breast cancer detection. Finally, we have succeeded in obtaining funding for related Projects, with the two breast pathology quality assurance studies, the American Cancer Society Study, the NCI funded hormone replacement therapy study, the NCI funded strategic studies in breast cancer surveillance study, and the AHRQ funded study on understanding variability in community mammography. An additional study has also been submitted to NCI, which proposes to explore breast cancer detection (both clinical and longer term outcomes) in women aged 70 and older.

The NHMN database is now an important resource for research on factors related to breast cancer risk, factors associated with screening behavior, and processes involved in breast cancer diagnosis. As the number of mammographic events per woman and breast pathology outcomes increases, we plan to expand our follow-up data collection to determine outcomes of women whose subsequent breast care was not recorded in our database, and to determine what motivated women with a false negative mammogram to seek care. We are confident that the NHMN database will continue to be an important resource for studies on patterns of care and accuracy in mammography in the coming years.

## LIST OF PERSONNEL RECEIVING PAY FROM THIS RESEARCH EFFORT

E. Robert Greenberg, MD
Patricia A. Carney, PhD
Marguerite Stevens, PhD
Steven P. Poplack, MD
Wendy Wells, MD
Anna Tosteson, ScD
Brenda Berube
Deirdre O'Mahony
Karen Burgess
Beth Harwood
Perry Spiegal
Mitzie Robinson
Keith Hamilton
Julie Wade
Martha Goodrich
M. Scottie Eliassen
Jeanette Ball
Judy
Scott Savioli
Bridgett Luce
Andrea Luce
Richard Ackerboom

## REFERENCES


1. New Hampshire Office of State Planning (September, 1994). 1994 Population Estimates. Concord, NH.

2. New Hampshire State Cancer Registry, 1993.

3. Chen VW, Wu XC. *Cancer in North America, 1990-1994. Volume II: Mortality.* Sacramento, CA: North American Association of Central Cancer Registries, 1998.

4. 1991 NH Behavioral Risk Factor Survey Report. NH Dept. of Health and Human Services, Concord, NH.

5. Azzopardi, J. Problems in Breast Pathology. (ed.). London: WB Saunders, Co, 1979.

6. Gillis, D., Dockerly, M. and Clagett, O. Preinvasive intraductal carcinoma of the breast. Surg Gynecol Obstet, 110: 555-562, 1960.

7. Holland, R., Peterse, J., Millis, R., Eusebi, V., Faverly, D., van de Vijver, M. and Zafrani, B. Ductal Carcinoma in situ: A proposal for a new classification. Seminars of Diagnostic Pathology, 11(3 (August)): 167-180, 1994.

8. Lampejo, O., Barnes, D., Smith, P. and Millis, R. Evaluation of infiltrating ductal carcinomas with a DCIS component: correlation of the histologic type of the in situ component with grade of the infiltrating component. Seminars in Diagnostic Pathology, 11(3 (August): 215-222, 1994.

9. Zafrani, B., Leroyer, A., Fourquet, A., Laurent, M., Trophilme, D., Validire, P. and Sastre-Garau, X. Mammographically-detected Ductal in situ carcinoma of the breast analyzed with a new classification. A study of 127 cases: correlation with estrogen and progesterone receptors, p53 and c-erb-B2 proteins and proliferative activity. Seminars of Diagnostic Pathology, 11(3 (August)): 208-214, 1994.

10. Kopans DB, D'Orsi CJ, Adler DED, et al. Breast imaging reporting and data system, 3rd ed. Reston, VA: American College of Radiology, 1998.

11. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro. Report of the International Workshop on Screening for Breast Cancer. JNCI, 1993;85(20):1644-1656.

12. Prochaska JO, DiClementa CC. Toward a comprehensive model of change. In WR Miller & N Healthier (eds)., Treating Addictive Behaviors: Process of Change (pp. 3-27). New York Plenum, 1986.

13. Colditz GA, Hankinson SE, Hunter DJ. The use of estrogens and progestins and

the risk of breast cancer in postmenopausal women. NEJM, 1995;332:1589-1593.

14. Newcomb PA, Longnecker MP, Storer BE, Mittendorf R, Baron J, Clapp RW, Bodgan G, Willett W. Long-term hormone replacement therapy and risk of postmenopausal breast cancer. Am J Epidemiol, 1995;142:788-795.

## LIST OF APPENDICES

# APPENDIX A – NHMN REPORT HANDLING POLICY

New Hampshire Mammography Network
NCCC • Evergreen Center, 46 Centerra Parkway, Suite 105, Lebanon, NH 03766-9907
Phone: 603-650-3414 Fax: 603-650-3415

# New Hampshire Mammography Network (NHMN) Report Distribution, Handling, and Data Alteration Policy (Revised August 10, 2000)

## Introduction

Each physician and facility contributing data to the New Hampshire Mammography Network (NHMN) may receive reports on the mammographic encounters they have provided to the Network. Outlined in this document are the policies for report handling, report development, and data alterations. Two sets of reports will be generated. They must be handled as outlined below.

**Level 1 Research Reports (Clinical Summary Reports)** will be provided to participating radiologists and mammography facilities for clinical application. Level 1 reports will contain patient-level information, including biopsy recommendations and outcomes.
**Level 2 Research Reports (Radiology Performance Reports)** will be generated for participating radiologists **only.** Level 2 reports will contain mammographic performance data (e.g. sensitivity, specificity, and positive predictive value) with comparisons to the state aggregate.

**Both Level 1 and Level 2 Research Reports must be handled in accordance with this policy.**

Level 1 and Level 2 Research Reports must be handled with the strictest confidentiality possible (in accordance with Institutional Review Board requirements for the protection of human subjects). Federal and State mechanisms exist, which protect the confidentiality of the data. However, these legislative acts will only protect our participants and databases from disclosure and litigation if these reports are handled appropriately. One breach by mishandling a report could threaten the protection we now have. This report handling policy has been developed with these issues in mind. **Level 1 & Level 2 Research Reports generated after January 9th, 1998 will NOT be distributed to individuals who have not signed this Report Distribution and Handling Policy (see Page 3).**

## Internal Report Development and Handling

**LEVEL 1 RESEARCH REPORTS will** contain clinically useful descriptive information. This report will allow facilities to track mammographic volumes, abnormal mammograms for which short follow-up or biopsy was recommended and pathology outcomes. For those facilities that choose to supply anonymous data on non-consenting patients this report will include information on this subset of patients as well as consenting women. **Because patient names are included in this report, it must be handled as confidentially as any medical record.** Dummy codes will be generated for NHMN on-site handling. A two step process will be used to produce these descriptive reports. One NHMN staff member will generate them and a second will place them in specially coded envelopes, which will then be added to the appropriate envelope for the radiologist to which the mailing will be sent.

**LEVEL 2 RESEARCH REPORTS** generated by NHMN staff will contain performance data and therefore will NOT identify the radiologists. Dummy codes will be generated for NHMN on-site handling. We will use a two step process for generating reports, where two different individuals are responsible for report generation and on-site handling. One person will be kept blind to the dummy code, but will have access to the database for report production and the other will be kept blind to the data source, but will apply the dummy code for processing and ultimate mailing.

For all **Level 2 Reports** that include comparative data, all radiologist specific data will be reported in sufficient aggregate to minimize the risk of identifying individuals or radiology practices, unless otherwise requested from the facility or radiology practice. (**ALL** radiologists in the group must agree to receive data with small cell sizes if this information is to be included).

**Report Handling by Participating Radiologists and Facilities**

Reports will be generated at six month intervals. They will be delivered to a radiologist designee at each mammography facility sent by express or certified mail. Allowable uses of reports include:

- **LEVEL 1 RESEARCH REPORTS (Clinical Summary Reports) ****

These reports are designed to facilitate practice management and patient tracking. They may be kept on file at mammography facilities according to the radiologist and facility's wishes.

- **LEVEL 2 RESEARCH REPORTS (Radiology Performance Reports)****

These reports identify Radiologist Read Groups (Practices) and Radiologists within a Read Group (Practice) and are identifiable sources of performance outcome measures. **These reports must be handled VERY CAREFULLY. They are ONLY to be reviewed by the individual (s) or groups who receive them. They will be provided to one individual radiologist at each practice who will be responsible for its handling and must be returned to the NHMN Project Office.**

**All reports contain only data that has been provided to the NHMN, which may not represent a complete picture of a facility or radiologist practice. The quality of data collection at facilities is critical for report accuracy.**

## ** LEVEL 1 & LEVEL 2 REPORTS SHALL *NOT* BE DUPLICATED

**We are currently protecting the database from discovery from potential litigation or other forced disclosure with a NH State Statute authorized by the NH State Health Commissioner and a Federal Certificate of Confidentiality. This protection is afforded because the database is a RESEARCH database. If data are used for non-research purposes or are handled inappropriately (as outlined above) this may threaten the protection now afforded.**

After your Level 2 Reports have been reviewed by all radiologist participants, we ask that you return them to the NHMN office in the self-addressed postage-paid mailer. This avenue of return will provide a receipt verifying that the NHMN has received the returned reports. We will shred the paper reports once they have been returned to our office. We will keep a computer disk that contains reports in a safety deposit box off-site. The safety deposit box will only be accessed after a request for access has been accepted by a majority of the advisory committee (of community radiologists). Access will be limited to a single designated NHMN staff member following authorization by a community radiologist representative of the NHMN Steering Committee. Newly generated reports will be shared only with the individual making the request.

**We ask that you NOT make photocopies, as this may pose a disclosure risk.**

## Inappropriate uses of reports include but are not limited to:

- Any media or marketing campaigns that use NHMN data for advertising, recruitment of patients, or other avenues of public information.
- Any sharing of reports with individuals not related to your professional practice or facility administration. (Level 2 reports should only be viewed by participating radiologists).
- Use of Level 2 data to satisfy professional credentialing.

### Data Alteration Policy

It is the goal of the NHMN registry staff to provide you with the most accurate reports possible. Because of patient consent issues, not every mammogram performed at your institution will be included in your report. We will do our utmost to generate accurate data on clinical performance. We understand that errors in data entry or administrative handling issues may occur on rare occasions, and thus have developed a policy on data alteration:

Data submitted to the database will be altered after a report has been generated ONLY if the facility or radiologist/pathologist can illustrate, using clear documentation, that an entry or other administrative error was made.

Agreement Statement

I have read and understand the contents of the New Hampshire Mammography Network (NHMN) Report Distribution and Handling Policy. By signing below, I agree to handle NHMN data reports as outlined.

| Radiologist Name | Signature | Signed Initials | Date | Witness |
|---|---|---|---|---|
| (INSERT NAME OF RAD) | _____ | _____ | _/__ /00 | _____ |
| New Radiologist(s) | | | | |
| _____ | _____ | _____ | __/__ /00 | _____ |
| _____ | _____ | _____ | __/__ /00 | _____ |
| _____ | _____ | _____ | __/__ /00 | _____ |
| _____ | _____ | _____ | __/__ /00 | _____ |

# APPENDIX B – PUBLICATION ON THE CHARACTERISTICS OF THE SCREENED POPLULATION IN NH

Journal of Community Health, Vol. 25, No. 3, 2000

# MAMMOGRAPHY IN NEW HAMPSHIRE: CHARACTERISTICS OF THE WOMEN AND THE EXAMS THEY RECEIVE

Patricia A. Carney, PhD; Martha E. Goodrich; Deirdre M. O'Mahony; Anna N. Tosteson, ScD; M. Scottie Eliassen, MS; Steven P. Poplack, MD; Steven Birnbaum, MD; Beth G. Harwood, MEd; Karen A. Burgess, BS; Brenda T. Berube, BS; Wendy S. Wells, MD; Jeanette P. Ball; Marguerite M. Stevens, PhD

ABSTRACT: New Hampshire (NH) is one of two states that has developed a population-based mammography registry. The purpose of this paper is to describe what we have learned about mammography use in New Hampshire. After collecting data for 20 months, the database contains almost 110,000 mammographic encounters representing 101,679 NH women, who range in age from 18 to 97 with a mean of 56.7 years (SD=10.91). Education levels are high with 92% having a high school education and 59% with some college. Forty-six percent report their primary insurance is private, 29% report HMO/PPO coverage, and 25% receive federal health care assistance. Risk factors represented in the database include (categories not mutually exclusive) advancing age (60% over age 50), hormone replacement therapy use by menopausal women (40.6%), and a family history of breast cancer (29%). Penetration of mammography relative to the NH population is higher for younger age groups (40–48% for those aged 44–64) than older age groups (34–39%

Patricia A. Carney is Assistant Professor of Community and Family Medicine, Martha Goodrich is Project Director for Breast Cancer Surveillance Research at the Norris Cotton Cancer Center, Deirdre O'Mahony is Database Manager for the New Hampshire Mammography Network (NHMN), Anna Tosteson is Associate Professor of the Departments of Medicine and Community and Family Medicine, M. Scottie Eliassen is Breast Pathology Coordinator for the NHMN, Steven P. Poplack is an Assistant Professor in the Department of Radiology, Steven Birnbaum is a Community Radiologist based in Parkland Medical Center in Derry, NH, Beth G. Harwood and Karen A. Burgess are Mammography Facility Field Coordinators for the NHMN, Brenda T. Berube is Director of Community Outreach for the Norris Cotton Cancer Center, Wendy Wells is Assistant Professor of Pathology, Jeanette P. Ball is Data Verifier for the NHMN, and Marguerite Stevens is Associate Professor of Community and Family Medicine; all, except Dr. Birnbaum, are at Dartmouth Medical School, Hanover, NH.

Requests for reprints should be addressed to: Patricia A. Carney, PhD, Dartmouth Medical School, 1 Medical Center Dr. HB 7925, Lebanon, NH 03756.

This research was funded by the U.S. Department of Defense (DASD17-94-J-4109).

We extend our sincerest appreciation to the community representatives of our Steering Committee, including Maureen Barrington, RTR (M), James Bronson, MD, Lynn Farnham, MD, David Hazemen, MD, Karen Jensen, MD, Laurie Ladd, RTR (M), Tim McCormack, MD, Lois Moore, MD, Paul Pullen, MD, and to the other radiologists, mammographers, and pathologists in New Hampshire whose commitment to this research project has made it possible.



0094-5145/00/0600-0183$18.00/0 

for those aged 65–84). The majority of mammographic encounters are routine screening exams (86%), often interpreted as negative or normal with benign findings (88%). Use of comparison films to interpret either diagnostic or screening mammography occurred in 86% of encounters. We have matched 3,877 breast pathology records to these mammographic encounters. The distribution of pathology outcomes for diagnostic exams was very similar to that for screening exams (approximately 65% benign, 17% invasive breast cancer, and 6% noninvasive breast cancer). Overall, we have designed a system that is well accepted by the NH community. Challenges include careful monitoring of data for coding errors, and a limitation of linking variables in mammography and pathology data. Data represented in this registry are a critical resource for research in mammographic screening and breast cancer early detection.

KEY WORDS: mammography screening; breast cancer detection.

## INTRODUCTION

Breast cancer is the most common cancer in women in the United States.[1] An estimated 184,300 new cases of invasive breast cancer were diagnosed in 1996,[1] and breast cancer incidence rates increased by 4% annually between 1982 and 1987.[1] A striking increase in frequency of noninvasive breast cancer, especially ductal carcinoma in situ, has recently been noted.[2,3] Much of the increased diagnosis appears to result from greater use of mammography and more frequent biopsy of suspicious findings. Almost 45,000 deaths (44,560) occurred from breast cancer in the US in 1996.[3] Several studies show that mammography screening could reduce breast cancer mortality by as much as 30% in women age 50 and over.[4-11] Recent research suggests that mammography in women between age 40 and 50 could also lead to reductions of about 20% in breast cancer mortality.[12] While health maintenance organizations can monitor screening processes and outcomes in their populations fairly readily, US public health departments are often not able to conduct such monitoring.

New Hampshire is one of two states funded to develop population-based mammography registries. Funding was received, in part, because breast cancer is especially problematic in New Hampshire (NH). It is the leading cancer in New Hampshire women with over 800 cases per year, representing 33% of female cancers statewide.[13] The mortality rate is 30.4 per 100,000, which is higher than the national rate of 24.1 per 100,000.[14] Early detection through population-based screening remains the best hope of reducing breast cancer morbidity and mortality. Women between the ages of 50 and 74 represent about 14% of New Hampshire's population.[15]

After initial design, development and pilot testing of a population-based mammography registry for the state of NH, which is described in detail elsewhere,[16] data collection began on May 1, 1996. Data represented in this registry are a critical resource for research in mammographic screening and breast cancer early detection. The purpose of this paper is to describe what we have learned about mammography use in New Hampshire, including the characteristics of the screened population, and the types and outcomes of imaging being performed. Comparing this information with data from New Hampshire Vital Statistics and the New Hampshire State Cancer Registry helps us determine who may not be receiving mammography screening or breast cancer early detection services in NH and the accuracy of data collection and linkage methods.

## METHODS

### The New Hampshire Mammography Network (NHMN)

The New Hampshire Mammography Network(NHMN) is a member of the National Cancer Institute sponsored Breast Cancer Surveillance Consortium (BCSC),[17] which is a network of eight states with mammography registries who submit data electronically to a statistical coordinating center for pooled analyses. All women who have a mammogram in a participating NH mammography facility are eligible to enroll in the NHMN. NHMN enrollment entails consenting to: provide data to the registry, allow tracking of mammographic information, release of medical and pathology information for linkage to radiology data, and future contact for research purposes.

The NHMN data collection instruments include the General Information Form (completed by patients), the Patient Intake Form (completed by the technologists) and the Radiologist Interpretation Form (completed by radiologists). The patients provide demographic and some risk factor information, such as age at menarche and first live birth. Registered mammography technologists query patients about presence of breast concerns; personal history of breast cancer; family history of breast cancer; including number of first degree relatives with breast cancer; current menopausal status (peri-menopausal or menopausal); whether the menopause was a natural part of aging or whether it was chemically or surgically induced; and whether the woman is currently taking hormone replacement therapy.

For each participant, radiologists record the indication for the exam, level of breast density, whether comparison films were used to interpret the exam, the assessment of the exam using the American College of Radiology Lexicon for mammogram assessment,[18] and future recommendations. Through prior agreement, NH pathology laboratories routinely send breast tissue reports to our study office, which are abstracted and entered into a separate pathology database.

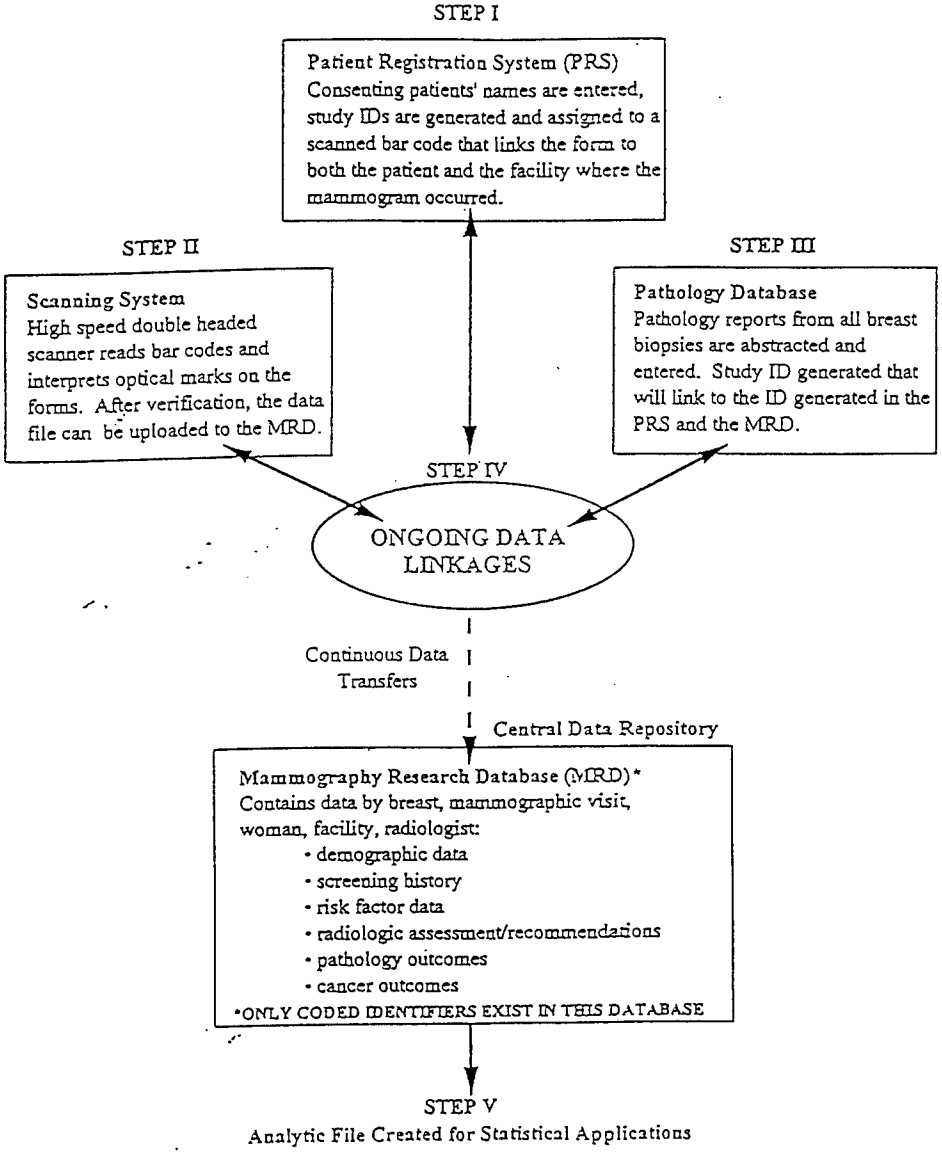
Oversight of the NHMN is accomplished through a steering committee made up of members of the research team, including a health services researcher, epidemiologist, radiologist, pathologist, biostatistician, and community representatives, including eight community radiologists, two registered mammography technologists, and one pathologist. The NHMN Steering Committee meets every six months to consider the design, development and approval of semiannual report formats that assist facility administrators and radiologists in monitoring patient outcomes and mammographic performance. The committee also reviews and approves proposed publications or any research proposals that involve use of NHMN data.

Data management activities take place at the NHMN Project office using NHMN data management systems. Secured access to computers, databases and network domains is maintained through using an isolated EtherNet Local Area Network (LAN) incorporating standard user identification and password authorization. Daily and weekly automated backup procedures are performed, with off-site safety deposit box storage of backup media.

The two major elements of the NHMN data collection systems are: 1) a group of relational databases for tracking women (Patient Registration System), mammographic encounters (Mammography Research Database), breast procedures and pathology outcomes (Pathology Database); and 2) a high-performance scanning system. The Patient Registration System (PRS) is the basic database for handling forms submitted by NHMN participating mammography sites. It is a registry of all enrollees in the NHMN Project, and contains identifying information such as a unique study identifier, name, address, date of birth, social security number, and maiden name, as well as information about the date and location of each mammographic encounter. This information is used to accurately link information from any participating mammography facility to the correct patient. Addresses are checked and updated each time a new mammographic encounter is entered for a woman. This database is used for mammography and pathology linkages to obtain our performance data for study outcome measures. Figure 1 illustrates NHMN databases and data flow.

FIGURE 1

NH mammography network databases and data flow.



The use of optical scanning allows for a streamlined survey format and rapid availability of results. Current NHMN surveys are double-sided, and the duplex scanner (which scans front and back sides of the page simultaneously) can process up to 2,500 forms per day. The data on the forms are interpreted using a combination of optical mark recognition (bubbles) and optical character recognition (text fields). Checks for range and consistency on individual fields have been programmed into the scanning system. All forms with ambiguities are visually verified by the scanner operator who makes any needed corrections.

When scanning is completed, the resulting files are uploaded from the scanning system to the Mammography Research Database (MRD), which is the repository for all of the NHMN mammogram data, maintained separately from the patient registration database to preserve confidentiality. The only link between the PRS and MRD systems is a unique study identifier. The MRD contains demographic data, screening history, medical history, radiologic assessments and recommendations. As data are continuously submitted to the NHMN, the expected total volume of mammographic events in the database will be approximately 300,000 by January 2000, a 65% increase from the current volume of mammographic encounters. We anticipate that approximately 70% of these will be repeat mammograms.

An important goal is to establish reliable record linkages between individuals enrolled in the Patient Registration System, the Pathology database and the New Hampshire State Cancer Registry. Accurate record linkage is dependent on the number of identifying variables present in the files being compared. The Pathology identifiers we receive are the patient's last, middle and first name and date of birth. The State Cancer Registry provides these identifying variables and in addition the Social Security number, street and zip code. With record matching, the higher the number of identifying variables, the lower the chances of there being a chance agreement. Although initially we relied on our own in-house matching programs, we needed a probabilistic method which would allow for conditions of uncertainty, since each field is subject to errors such as dates transposed, missing values, or incorrect spelling. To address this, we recently implemented a software application that uses the methodology originally developed and tested at the US Census Bureau for census under count estimation.[19] With this software in place, the reliability of the NHMN matching methods is statistically justifiable and reproducible. In addition, we estimate that accuracy in linkages between mammography and breast pathology has increased by about 1% with probabilistic matching methods compared to our in-house matching program.

## Defining Variables

Data presented in this analysis were taken from the start date (May 1, 1996) through December 31, 1997. After collecting data for 20 months, data queries were made from our database systems to describe characteristics of the NHMN participants and imaging exams they received. Screening mammograms were defined as bilateral two view (craniocaudal and mediolateral) mammograms which were not influenced by concerns about symptoms, positive clinical breast examinations or a mammogram within the previous 270 days. Though data are collected by breast, mammograms were classified by the most serious interpretation, and breast density is reported by densest breast. Non-residents of New Hampshire are collected in our database (n = 3,876 for this time period), though these women were excluded from these analyses. To assess mammographic penetration in NH, we obtained population statistics by age category from the NH Office of State Planning and divided these estimates by the same age categories represented in our registry. This gives us a very rough estimate of mammography use by NH women. The number of mammograms performed on NH women out of state is not known.

# RESULTS

To date, we have obtained written consent from over 87% of women who have entered one of our 41 participating mammography facilities. Our rate of participation by mammography facilities in NH is 91%, which, based on volume estimates for nonparticipating sites (adjusted for non-resident status), indicates our capture is an estimated 90% of women getting mammography in NH since May 1st, 1996 (nonparticipating facilities are low volume centers). Our database currently contains data on 118,549 mammographic encounters representing 101,679 consenting NH women and 8,751 anonymous encounters (women who did not consent to have their identifying data included in the registry). We collect encounter data on both consenting and nonconsenting women; however linkages to subsequent mammography and/or breast pathology are not possible for nonconsenting women. The anonymous encounters are useful for comparing characteristics of consenting and nonconsenting women.

Women represented in the registry (n = 101,679) range in age from 18–97 with a mean age of 56.7 (SD = 10.91) years. Approximately 6% are under age 40; 30% are 40–50; 27% are 51–60; 18% are 61–70; 12% are 71–80; and 4% are over age 80. The majority of women in the registry are

Caucasian (98.5%). The remaining ethnic backgrounds represented include 0.4% Asian, 0.3%

African American, 0.5% Native American, and 0.3% Hispanic. We found the educational status of participants to be relatively high with 33% high school graduates, 30% with some college, 16% college graduates and 13% with a post graduate education. Eight percent were not high school graduates. Almost 46% of women in the registry report having private insurance, 29% report HMO or PPO coverage, 20% report having Medicare coverage and 3% report having Medicaid coverage, 3% report having no health insurance, 2% report having CHAMPUS or CHAMPVA and 7% report other coverage.

Table 1 provides a breast cancer risk factor profile of women represented in the NHMN. Advancing age is the most common risk factor, though we note a relatively high use of hormone replacement therapy in women who are no longer menstruating (41%) and a high percentage of women with a family history of breast cancer (29%) (including mother [38%], sister[28%], daughter [5%], and other family members [34%]). Women with a personal history of breast cancer (5%) have a low representation in the data, and women with an extended menopause (women over age 50 with periods) have a less than 1% representation.

Table 2 compares the age representation of women in the NHMN database to that of the general population. Based on these estimates, the

TABLE 1

Breast Cancer Risk Factor Profile of Women Represented in the New Hampshire Mammography Network

| *Risk Factor* | *% Represented in Database (n=101,679)* |
|---|---|
| Age 50 and Over | 60.6 |
| Hormone Replacement Therapy Users* | 40.6 |
| Family History of Breast Cancer | 29.1 |
| Age at Menarche < 12 | 19.6 |
| Age at First Live Birth >30† | 7.0 |
| Personal History of Breast Cancer† | 5.0 |
| Women Over 50 With Periods | 0.7 |

* In women who are no longer menstruating (n = 70,551)
† In women over age 30 (n= 98,222)

TABLE 2

Estimated Penetration of Mammography in New Hampshire by Age Category

| *Age Category** | *New Hampshire Mammography Network Number (%)* | *New Hampshire Population† Number (%)* | *Estimated % Penetration* |
|---|---|---|---|
| 35–44 | 19,436 (19.4) | 118,250 (36.8) | 16.4 |
| 45–54 | 32,238 (32.0) | 80,300 (25.0) | 40.1 |
| 55–64 | 22,041 (21.9) | 46,200 (14.4) | 47.7 |
| 65–74 | 16,561 (16.4) | 41,800 (13.0) | 39.6 |
| 75–84 | 8,764 (8.7) | 25,850 (8.0) | 33.9 |
| >85 | 1,575 (1.6) | 8,800 (2.7) | 17.9 |
| Total | 100,615 (100) | 321,200 (100) | |

* Age categories are fixed in NH vital statistics report, making comparisons to age category based on mammography screening recommendations impossible.
† based on NH vital statistics.

penetration of mammography in NH women between the age categories of 45–54 and 55–64 years is approximately 40 and 48% respectively. After age 64, mammographic penetration drops as age categories advance. We found 39% penetration in the 65–74 age category, and 34% in the 75–84 age category.

NH vital statistics also indicate that 10.2% of NH women over age 65 have incomes below the federally designated poverty level. Although income levels are not collected in the NHMN database, it is interesting to note that 12.4% of the women over 65 have not received a high school diploma.

A recent publication describes the participating mammography facilities in detail.[16] Briefly, of the 41 facilities contributing data to the registry, 54% represent hospital-based facilities, 22% are clinic-based affiliates of the hospitals represented, 11% are in radiologists' private offices, 9% are in non radiologists' offices, 2% are in women's health centers.

Table 3 outlines the imaging services used to determine the American College of Radiology assessment categories[18] for 109,798 mammograms (those of consenting women only) represented in the NHMN database by indication for exam (screening versus diagnostic). Screening mammograms make up almost 86% of the mammograms in the database,

TABLE 3

Imaging Services Used to Determine American College of Radiology Assessment Categories for Mammograms Represented in the New Hampshire Mammography Network

| *Imaging/Assessment* | *Screening Exam (86%) n=94,533* | *Diagnostic Exam (14%) n=15,265* |
|---|---|---|
| IMAGING | | |
| • Standard 2 View Mammography (Craniocaudal and Mediolateral) Only Used | 94.1 | 86.2 |
| • Standard 2 View Mammography Plus Supplemental Ultrasound Used | 6.2 | 11.8 |
| • Standard 2 View Mammography Plus Supplemental Additional Views Used | 9.6 | 20.6 |
| Use of Comparison Films (for interpretation only) | 87.5 | 86.9 |
| ACR ASSESSMENT | | |
| • ACR 0—Needs Additional Assessment | 2.9 | 4.2 |
| • ACR 1—Negative | 78.2 | 43.0 |
| • ACR 2—Benign Finding | 9.9 | 19.5 |
| • ACR 3—Probably Benign | 6.8 | 26.2 |
| • ACR 4—Suspicious Abnormality | 1.5 | 5.9 |
| • ACR 5—Highly Suggestive of Malignancy | 0.6 | 1.2 |

and the vast majority (88.1%) are assessed as negative or normal with benign findings. Even in diagnostic mammography, suspicious or highly suggestive of malignancy categories are very seldom used (less than 8% of diagnostic mammograms).

To date we have matched 3,788 breast pathology reports to mammographic encounters recorded in the database. In matching pathology, we link the pathology report to the first mammogram performed within 365 days, which allows us to identify the event that initiated pathology follow-up (presentation for a screening exam or diagnostic exam). Of the pathology matches, 82% matched to initial screening mammograms and

18% matched to initial diagnostic mammograms. Of the pathology reports that matched to screening exams, 64% were benign, 20% were invasive, and 6% were noninvasive (approximately 92% ductal carcinoma in-situ, 8% lobular carcinoma in-situ excluded). The remaining were atypical (6%), unsatisfactory (4%) or suspicious (< 1%). We found the distribution of pathology outcomes for diagnostic exams to be very similar to that of screening exams. Of the pathology that matched to diagnostic exams, 67% were benign, 17% were invasive, 6% were noninvasive, and the remaining reports were atypical (6%), unsatisfactory 4% or suspicious (< 1%).

## DISCUSSION

Breast cancer is a significant problem in NH. In developing and implementing our state-wide mammography network, we have found that community acceptance of the project by NH women is very high (an estimated 90%). This consent rate is a credit not only to the participating women themselves, but also to the mammography technologists who introduce the women to the project, and the radiologists who are committed to optimizing mammography in NH. Maintaining positive relationships and rapport with technologists and radiologists is a critical factor in the success of our registry.[20]

Our goal was to design a data collection process that was simple and easily incorporated into the mammography appointment. We have found that participation in paper-based data collection activities by mammography facilities, their staffs, and radiologists is high. This is facilitated by the fact that the reports we generate for facilities and radiologists assist in monitoring patients' outcomes as well as mammography performance. We are currently testing a computer-based data collection system, which will be used to collect data electronically. This system will allow facilities to automatically generate reports and reminder letters as well as allow us to continue collecting high quality standardized data.

We found that the screened population in NH were well educated with over 59% having some college and only 8% without a high school education. This likely reflects income status as also being relatively high. We also found women in the database were well insured with only 3% with no health insurance. We found that managed care penetration in this population was fairly small with fee for service being the most common form of insurance coverage. The ethnic distribution matches that found in our population, being predominantly Caucasian with an age distribu-

tion (over 60% over age 50) that matches indications for screening and diagnostic mammography.

Our data indicate that 41% of peri-or post menopausal women are current hormone replacement therapy (HRT) users. Prevalence estimates of the use of HRT in the general US population range from 5.3%[21] to 39.3%.[22,23] Indications for HRT use include relief of post menopausal symptoms, such as hot flashes,[24] and reductions in rise for conditions such as osteoporosis,[25,26] urogenital atrophy[27] and cardiovascular disease (coronary artery disease and stroke).[28,29] Several studies, including a large pooled analysis,[30] have indicated that HRT is associated with increased risk of breast cancer,[31-37] We would, therefore, expect to see that HRT use is higher in women being screened than in the general population, despite the fact that this finding remains controversial.[38,39]

HRT use is also associated with increased density in breast tissue.[40-42] Breast density is associated with a four-to sixfold increase in breast cancer risk and less accurate mammography.[43,44] Our data indicate the density distribution of screening mammograms in NH is 15% fat, 46% scattered, 31% heterogeneously dense, and 8% extremely dense. How these rates compare to breast density in other mammography registries with similar use of HRT is unknown, but is an area for further research in light of the controversial relationships between HRT and breast cancer risk.

We also learned that mammography penetration in NH is highest in women aged 55–64. However, it is still less than half of the eligible population. Penetration in women over 65 is less than 40%. One limitation in this analysis is that our NHMN data are based on 20 months of data collection, which may be an underestimate of mammography use, especially in women under age 50. As data are collected over time, we expect these estimates will become more accurate. In any case, mammography has not reached its full potential in NH. While health maintenance organizations can easily monitor the screening status of their members, the ability of public health officials to monitor mammography screening is limited. By comparing information in our database to other data available from NH vital statistics, we can track estimates for mammographic penetration in NH women as well as cancer outcomes on an ongoing basis. This information will inform public health programs that target difficult to reach women, such as the Centers for Disease Control and Prevention-funded Breast and Cervical Cancer Early Detection Program.

The distribution of ACR assessment categories by screening versus diagnostic exam revealed no surprises. We expected that the majority of screening mammograms would be interpreted as negative or normal with benign finding and over 88% were so interpreted. Over 60% of the diag-

nostic exams were also determined to be normal, with an additional 26% interpreted as probably benign. The diagnostic exams yielded an almost fourfold increase in suspicious abnormalities, and a twofold increase in those highly suggestive of malignancy. We also found the pathology yield by diagnostic type in screening mammograms was very similar to diagnostic mammograms. Despite the absence of clinical findings, a positive finding on a screening mammogram does not predict any particular pathology outcome but merely represents the same possible diagnostic differential as when a clinical finding is present and precedes a diagnostic mammogram, which underscores the importance of mammography screening.

By linking to pathology outcomes of women in NH, we will understand the performance of community-based mammography (sensitivity, specificity, positive and negative predictive value). Once enough time has elapsed for all participating facilities to have contributed adequate data to provide stable rates of performance, and sufficient numbers of screening cases can be closed out (no breast pathology linkages for 365 days), we can calculate these measures.

The NHMN registry is an important public health resource. We are able to monitor mortality, stage, and other prognostic factors of disease related to the natural history of breast cancer, as well as what risk factors are related to incidence and what cascade of imaging care results in the best breast cancer detection outcomes. We have also sent data on over 80,000 encounters to the National Cancer Institute funded Breast Cancer Surveillance Consortium Statistical Coordinating Center.

Limitations do exist in collecting and interpreting our data. Because of the very low representation of minorities in New Hampshire, our understanding of mammography in these groups is limited. In addition, it is our belief that a percentage of mammographic encounters of women residing in NH are performed in Massachusetts or other bordering states. The degree to which this occurs is currently unknown.

The ACR lexicon for mammographic assessment and recommendations is relatively new and the registry will allow for testing of the lexicon. We have found that coding errors do occur in data collection, which can result in a misclassification error. We have noted this when breast laterality is coded incorrectly. When this occurs, it can be very difficult to link mammography and pathology to the actual breast involved. When reports are generated for facilities, the data are reviewed. If coding errors have occurred, documentation of the error is sent to the project office. Appropriate changes are made in the database, and the changes are logged.

Pathology matching has also presented a challenge because there

are limited variables available for matching pathology and mammography identifiers We currently have name, date of birth, and date of exam. If Social Security numbers were available as a pathology identifier (as is in mammography) our matching might be enhanced. The recent implementation of a new record linkage application that provides a statistically justifiable methodology has increased precision and increased matching.

Despite the challenges that have arisen in collecting and interpreting the data for reports and research purposes, the ultimate benefits that women in our state could experience are many. Currently, our database indicates that the rate of adherence to screening mammography recommendations in women age 50 and older (excluding those with a previous history of breast cancer) is approximately 70%. We have just begun two related studies examining how anxiety and risk may influence annual mammography adherence in women in this age group. As part of these investigations, we will be testing interventions to improve interval screening in women age 50 and older.

Additional opportunities for research include: examining how hormone replacement therapy, which increases breast density, influences mammographic interpretation; and testing educational interventions for radiologists that would assist them in understanding the benefits of specific time intervals for follow-up and reducing the time period between an abnormal mammogram and a definitive diagnosis of cancer. These additional studies depend on ongoing support for the registry. Core registry operations cost approximately $1.00 per mammogram or about $130,000 annually. Though research funds can help defray these costs since infrastructure support is often difficult to obtain, additional support for this research is needed. We are currently pursuing corporate sponsorship in addition to research funding for continued operation of this important population-based public health resource.

## REFERENCES

1. *American Cancer Society. Cancer Facts and Figures–1996.* Atlanta, GA: American Cancer Society: Atlanta, GA. Publication number: 96-300M-No. 5008.96.
2. Ernster V, Barclay J, Kerlikowske K, Grandy D, Henderson C. Incidence of and treatment for ductal carcinoma in situ of the breast. *JAMA* 1996; 12:913–8.
3. Choi W, Parker B, Pierce J, Greenberg E. Regional differences in the incidence and treatment of carcinoma in situ of the breast. *Cancer Epidem Biomarkers Prev* 1996; 4:317–20.
4. Baker LH. Breast Cancer Detection Demonstration Project: Five-year summary report. *CA Cancer J Clin* 1982; 32:94–225.
5. Shapiro S, Venet W, Strax P, Venet L, Roeser R. Ten to fourteen-year effect of screening on breast cancer mortality. *J Natl Cancer Inst* 1982; 69:349–55.

6. Verbeek AL, Hendricks JH, Holland R, Mravunac M, Sturmans F, Day NE. Reduction of breast cancer mortality through mass screening with modern mammography. First results of the Nijmegen project, 1975–1981. *Lancet* 1984; 1:1222–4.
7. Tabar L, Fagerberg CJ, Gad A, Baldertorp L, Holmberg LH, Grontoft O. Reduction in mortality from breast cancer after mass screening with mammography. Randomized trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet* 1985; 1:829–32.
8. deKoning HJ, Francheboud J, Boer R, Verbeek AL, Collette HJ, Hendricks JH. Nation-wide breast cancer screening in The Netherlands: Support for breast cancer mortality reduction. *Int J Cancer* 1995; 60:777–80.
9. Smart CR, Hendrick RE, Rutledge JH, Smith RA. Benefit of mammography screening in women 40–49 years. *Cancer* 1995; 75:1619–26.
10. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro. Report of the International Workshop on Screening for breast cancer. *J Natl Cancer Inst* 1993; 85:1644–56.
11. Mettlin CJ, Smart CR. The Canadian National Breast Screening Study. An appraisal and implications for early detection policy. *Cancer; Diagnosis Treatment, Research* 1993; 72:1461–5.
12. Fletcher SW. Why question screening mammography for women in their forties? *Radiol Clin North Am* 1995; 33:1259–71.
13. New Hampshire State Cancer Registry, 1993.
14. Chen VW, Wu XC. *Cancer in North America, 1990–1994. Volume II: Mortality.* Sacramento, CA: North American Association of Central Cancer Registries, 1998.
15. New Hampshire Office of State Planning (September, 1994). 1994 Population Estimates. Concord, NH.
16. Carney PA, Poplack SP, Wells WA, Littenberg B. Development and Design of a Population-Based Mammography Registry: The New Hampshire Mammography Network. *Am J Roentgenol* 1996; 167:367–72.
17. Ballard-Barbash R, Taplin SH, Yankaskas BC, Ernster VL, Rosenberg RD, Carney PA, Barlow WE, Geller BM, Kerlikowske K, Edwards BK, Lynch CF, Urban N, Chrvala CA, Key CR, Poplack SP, Worden JK, Kessler LG. Breast Cancer Surveillance Consortium: A national mammography screening and outcomes database *Amf Roentgenol* 1997; 169:1001–8.
18. Kopans DB, D'Orsi CJ, Adler DED, et al. *Breast Imaging Reporting and Data System*, 3rd ed. Reston, VA: American College of Radiology, 1998.
19. Jaro MA. Probabilistic Linkage of Large Public Health Data Files. *Statistics in Medicine*, 1995; 14: 491–8.
20. Burgess, KA, Harwood B, Robinson M, Carney, PA. A Qualitative Study of Mammographers' Participation in a Large Database Research Study. *Radiologic Technology* 1999; 70(5):453–460.
21. Derby CA, Hume AL, Barbour MM. Correlates of postmenopausal estrogen use and trends through the 1980s in two southeastern New England communities. *Am J Epi*, 1993; 137:1125–35.
22. Criqui MH, Suarez L, Barrett-Connor E. Postmenopausal estrogen use and mortality. Results from a prospective study in a defined, homogenous community. *Am J Epi* 1975; 102:209–16.
23. Johannes CB, Crawford SL, Posner JG, McKinlay SM. Longitudinal patterns and correlates of hormone replacement therapy use in middle-aged women. *Am J Epi* 1994; 140:439–52.
24. Wood AJ. Hormonal treatment of postmenopausal women. *NEJM* 1994; 330:1062–71.
25. Session DR, Kelly AC, Jewelewicz R. Current concepts in estrogen replacement therapy in the menopause. *Fertil Steril* 1993; 59:227–84.
26. Felson DT, Zhang Y, Hannon MT, Keil DP, Wilson PW, Anderson JJ. The effect of postmenopausal estrogen therapy on bone density of elderly women. *NEJM* 1993; 273:199–208.
27. Fantl JA, Cardozo L, McClish DK. Estrogen therapy in the management of urinary incontinence in postmenopausal women: The first report of the hormones and urogenital therapy committee. *Obst Gynecol* 1994; 83:12–8.
28. The Writing Group for the PEPI Trial. Effects of estrogen or estrogen/progestin regimens on heart disease risk factors in postmenopausal women. The Postmenopausal Estrogen/Progestin Interventions (PEPI) Trial. *JAMA* 1995; 273:199–208.
29. Finucane FF, Madans JH, Bush TL, Wolf PH, Kleinman JC. Decrease risk of stroke among postmenopausal hormone users. *Arch of Int Med* 1993; 153:73–9.
30. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: Collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet* 1997, 350:1047–59.

31. Colditz GA, Hankinson SE, Hunter DJ. The use of estrogens and progestins and the risk of breast cancer in postmenopausal women. *NEJM* 1995;332:1589-93.
32. Helzlsouer KJ, Couzi R. Hormones and breast cancer. *Cancer* 1995; Suppl 76:2059-63.
33. Stanford JL, Weiss NS, Voigt LF, Daling JR, Habel LA, Rossing MA. Combined estrogen and progestin hormone replacement therapy in relation to risk of breast cancer in middle-aged women. *JAMA* 1995; 274:137-42.
34. Tzingounis V, Cardamakis E, Ginopoulos P, Argiropoulos G. Incidence of benign and malignant breast disorders in women taking hormones (contraceptive pill or hormone replacement therapy). *Ca Res* 1996; 16:3397-4000.
35. Smith HO, Kammerer-Doak DN, Barbo DM, Sarto GE. Hormone replacement therapy in the menopause: A pro opinion. *CA Cancer J Clin* 1996; 46:343-63.
36. DiSaia PJ, Grosen EA, Kurosaki T, Gildea M, Cowan B, Anton-Culver H. Hormone replacement therapy in breast cancer survivors: a cohort study. *Am J Obstet Gynecol* 1996; 176:1494-98.
37. Persson I, Thurfjell E, Bergstrom R, Holmberg L. Hormone replacement therapy and the risk of breast cancer. Nested case-control study in a cohort of Swedish women attending mammography screening. *Int J Ca* 1997; 72:758-61.
38. Breckwoldt M, KeckC, Karck U. Benefits and risks of homone replacement therapy. *J Steroid Biochem Molec Biol* 1995; 53:205-8.
39. Mayeaux EJ, Jr, Johnson C. Current concepts in postmenopausal hormone replacement therapy. *J Fam Pract* 1996; 43:69-75.
40. Laya MB, Gallagher JC, Schreiman JA. Effect of postmenopausal hormone replacement therapy on mammographic density and parenchymal pattern. *Radiology* 1995; 196:443-47.
41. Stomper PC, Van Voorhis BJ, Ravkinar VA. Mammographic changes associated with post menopausal hormone replacment therapy: A longitudinal study. *Radiology* 1990; 174:487-90.
42. Leung W, Goldberg F, Zee B, Sterns E. Mammographic density in women on postmenopausal hormone replacement therapy. *Surgery* 1997; 122:669-74.
43. Staftlas AF, Hoover RN, Brinton LA, Szklo M, Olson DR, Salane M. Mammographic densities and risk of breast cancer. *Cancer* 1991; 67:2833-8.
44. Boyd NF, Byng JW, Jong RA, Fishell EK, Little LE, Miller AB. Quantitative classification of mammographic densities and breast cancer risk: Results from the Canadian National Breast Screening Study. *JNCI* 1995; 87:670-5.

# APPENDIX C – PUBLICATIONS ON BREAST PATHOLOGY AGREEMENT STUDIES

# REPORTS

## Statewide Study of Diagnostic Agreement in Breast Pathology

*Wendy A. Wells, Patricia A. Carney, M. Scottie Eliassen, Anna N. Tosteson, E. Robert Greenberg**

***Background:*** **This study assessed the degree of diagnostic agreement among community-based general pathologists reading slides of representative breast tissue specimens and tested whether diagnostic variability is associated with type of breast specimen (e.g., core needle or excisional biopsy) or slide quality.** ***Methods:*** **Twenty-six of the 44 eligible pathologists working at community-based pathology practices in New Hampshire participated. Each pathologist evaluated slides of breast tissue obtained from 30 case subjects randomly selected from a statewide breast pathology database. The diagnostic categories used were benign, benign with atypia, noninvasive malignant, and invasive malignant. The levels of agreement (i.e., kappa coefficients) for the diagnoses were assessed.** ***Results:*** **Agreement was high among pathologists for assignment of diagnostic category (kappa coefficient = 0.71) and was nearly perfect for their selection of benign versus malignant categories (kappa coefficient = 0.95). There was less agreement for the categories of noninvasive malignant and benign with atypia (kappa coefficients of 0.59 and 0.22, respectively). There was no apparent relationship between levels of diagnostic agreement and specimen type or perceived slide quality.** ***Conclusions:*** **Diagnostic agreement for breast tissue specimens is high overall among community-based pathologists, but clinically relevant disagreements may occur in the assessment of noninvasive malignant diagnoses. The establishment of reread policies for certain diagnostic categories may reduce the possibility that diagnostic misclassification will lead to overtreatment or undertreatment. The high degree of diagnostic reproducibility for invasive cancerous lesions of the breast suggests that it is unnecessary for a central review of these lesions in national cancer trials. [J Natl Cancer Inst 1998;90:142–5]**

The frequency of diagnosis of breast cancer has increased markedly over the past 2 decades, particularly for noninvasive ductal carcinoma *in situ (1,2)*. Much of this increase results from greater use of high-quality mammography and more frequent biopsy of suspicious findings. Previous studies *(3,4)* have found relatively poor agreement among pathologists in their diagnostic assessments of breast disease, but these studies have largely used pathologists in academic centers with a special interest in breast pathology, and the slides reviewed were from cases with challenging histologic features. There is scant information on the reproducibility of diagnoses provided by community based pathologists *(5–7)*, and no data have been published from a representative mix of biopsy specimens interpreted by pathologists in the United States. This report describes the degree of interobserver agreement for breast diagnoses among community-based general pathologists in New Hampshire.

## Methods

The study was approved by an institutional committee for the protection of human subjects and endorsed by the New Hampshire Society of Pathologists. We sent recruitment letters and information detailing the proposed study and the lead investigator (W. A. Wells) met with each of the 44 eligible pathologists in New Hampshire. To be eligible to participate, a pathologist must have been actively practicing general surgical pathology in New Hampshire, have regularly evaluated breast tissue, and have reported no plans to retire or relocate within the study period. Each participant returned a signed consent form.

Forty-four pathologists met the criteria for eligibility, and 35 (80%) of these pathologists—representing 14 (82%) of the state's 17 hospitals with laboratories that process breast tissue specimens—agreed to submit breast pathology reports for all biopsied and excised breast tissue beginning in January 1996. Six pathologists from the only academic center in the state were also included. Data on specimen type (e.g., core biopsy or excisional biopsy) and diagnosis were entered into a central database. Pathologists also provided information on demographic/practice characteristics, usual content of breast pathology reports, and tissue processing methods.

After 3 months of data collection, the pathology database held information on 502 biopsy specimens. After stratifying the cases in the database by diagnosis, a random number table was used to select 30 case subjects with diagnoses representative of the distribution of all diagnoses in the database. We asked pathologists who had submitted the selected reports to submit four recut tissue slices of a representative block from the case. Each recut specimen, from a formalin-fixed, paraffin-embedded tissue block, was 4-μm thick and was stained with hematoxylin–eosin under standard conditions. The recut specimens were reviewed (by W. A. Wells) to ensure that the same histopathologic material was present on each recut tissue slice. The slides were masked and organized into four complete sets, each mailed according to a structured rotation schedule so that each pathologist read one set of 30 slides. Of the selected slides, nine were derived from image-guided core biopsy specimens (stereotactic or ultrasound guided) and 21 from excisional biopsy and mastectomy specimens.

All participating pathologists used a standard reporting sheet to record their interpretations of each slide in the circulated set. Summarized categories of diagnosis were: benign, benign with atypia, noninvasive malignant, and invasive malignant. The pathologists also evaluated each slide for processing, staining, and sectioning quality by categories of excellent, very good, satisfactory, and unsatisfactory. For slides with quality perceived to be less than very

*Affiliations of authors:* W. A. Wells (Department of Pathology), P. A. Carney, M. Scottie Eliassen, E. Robert Greenberg (Departments of Community and Family Medicine and Medicine, and The Norris Cotton Cancer Center), A. N. Tosteson (Department of Community and Family Medicine and The Center for Evaluative Clinical Sciences), Dartmouth Medical School, Hanover, NH.

*Correspondence to:* Wendy A. Wells, M.D., Department of Pathology HB 7999, Dartmouth-Hitchcock Medical Center, 1 Medical Center Drive, Lebanon, NH 03756. E-mail: Wendy.Wells@dartmouth.edu

*See* "Notes" following "References."

good, the participants were asked to detail the deficiency. Possibilities included inadequate tissue fixation, poor tissue processing (alcohol clearing and paraffin infiltration), section artifacts (thickness and wrinkles), and suboptimal staining. Participants were blinded to the original diagnosis and to each others' readings.

To assess diagnostic agreement, we computed a kappa statistic (i.e., coefficient) for the overall agreement in all four diagnostic categories and for comparisons between categories (e.g., benign cases versus malignant categories and noninvasive malignant cases versus all other categories). The kappa statistic estimates the level of agreement, after accounting for agreement that would be expected by chance alone. Kappa statistics less than 0.4 represent fair to poor agreement, values of 0.4 to 0.8 represent moderate to good agreement, and values over 0.8 represent excellent agreement *(8)*. The impact of slide quality and sample source was also examined in subgroup analyses.

Continuing Medical Education credits were awarded to all pathologists completing the project, and each was sent a report comparing his/her individual interpretations with the statewide aggregate results. The results were presented at the annual meeting of the New Hampshire Society of Pathologists.

## Results

Twenty-six (74%) of the 35 pathologists who submitted reports to the database took part in the slide review and contributed data to the current analyses. The characteristics of the 26 participants differed little from those of the 17 eligible nonparticipating pathologists (Table 1). Of the nine who did not provide data for the analyses, one (W. A. Wells) was ineligible (had viewed the slides during the selection process), three were excluded because they read the study slides as a group, and five chose not to participate in this portion of the project.

We received a total of 775 review diagnoses from the 26 participants who nearly all provided a diagnosis for each of the 30 slides. Five diagnosis review forms were left entirely blank, one each by five pathologists. The distribution of diagnoses for the study slides [489 (63%) benign, 47 (6%) benign with atypia, 66 (9%) noninvasive malignant, and 173 (22%) invasive malignant] was comparable to the distribution of diagnoses reported to the breast pathology database [330 (66%) benign, 18 (4%) benign with atypia, 28 (6%) noninvasive malignant, and 122 (24%) invasive malignant] at the time the random sample of 30 cases (representing 30 patients) was chosen.

There was a clear consensus on the diagnosis for almost every case, with complete agreement for 11 (37%) of the 30 cases (Table 2). For differentiation between benign and malignant categories, there was complete agreement for 22 (73%) of the cases. Clinically relevant diagnostic variations were observed in eight (27%) cases (N, O, P, Q, S, T, U, and V), with discrepancies in benign versus malignant diagnoses by one pathologist. For two of these cases (N and P), the majority diagnosis was benign with one diagnosis of invasive malignant. For three cases (X, Y, and Z), there was substantial disagreement between noninvasive malignant and invasive malignant. For six (20%) cases (H–M), the majority diagnosis was benign, but one pathologist made a diagnosis of benign with atypia. For these six cases, as well as for cases N, O, P, Q, S, T, U, and V, identification of the one pathologist who recorded a discordant diagnosis compared with all of the other pathologists revealed a different person in every case.

The kappa coefficient confirmed a high level of agreement for assignment of diagnostic category (kappa coefficient = 0.71) and near perfect agreement for the distinction between the two benign versus the two malignant categories (kappa coefficient = 0.95). Less reproducible diagnostic categories, compared with others, were the benign with atypia and noninvasive malignant, with kappa coefficients of 0.22 and 0.59, respectively (Table 3).

Only 30% of the participants indicated that they routinely review core biopsy specimens in their daily practice. However, the kappa coefficient for the nine image-guided core biopsy specimens was 0.85 overall and 0.98 for distinguishing between the benign and malignant categories. These figures were only slightly lower for the noncore biopsy specimens (0.60 and 0.85, respectively). Kappa coefficients for distinguishing between diagnoses of noninvasive cancer versus the other categories were 0.57 and 0.60 for the core and noncore specimens, respectively. The recognition of histologic special type invasive tumors (lobular and colloid) in both the core and noncore specimens was excellent.

For slides where reviewers rated the quality lower than very good, the most commonly cited deficiencies were fixation and staining quality. However, reduced quality did not seem to affect diagnostic agreement. The kappa coefficient for slides interpreted as of high quality (rated by ≥75% of participants as excellent, very good, or satisfactory) was 0.64. For slides classified as unsatisfactory or rated by greater than or equal to 25% of reviewers as only satisfactory, the kappa coefficient was 0.69. The twelve pathologists classifying 17 slides as unsatisfactory, attributed the poor quality roughly equally to fixation, staining, sectioning, and processing. No single laboratory was responsible for consistently substandard slide quality.

Nineteen (66%) of 29 pathologists completed our survey about breast pathology reread procedures (defined as a second pathologist giving an independent evaluation of all or some breast pathology cases). Of these, 16% reported rereading all breast tissue cases (benign and malignant). An additional 37% reported rereading all malignant, benign with atypia, and noninvasive malignant cases. Rereading of specimens originally diagnosed as benign with atypia or noninvasive malignant was reported for 21% and 26% of cases, respectively.

## Discussion

This study indicates a high level of diagnostic agreement for the type of breast pathology material routinely reviewed in practice by community pathologists in New Hampshire. None of these pathologists has a special expertise in breast pathology.

There were high levels of agreement (i.e., high kappa coefficients) for all four

**Table 1. Characteristics of eligible participating and nonparticipating pathologists***

| Characteristic | Eligible nonparticipants (n = 17) | Participants (n = 26) |
|---|---|---|
| Median age in y (range) | 53 (35–65) | 47 (36–65) |
| Median time in practice in y (range) | 15 (4–20) | 16 (2–37) |
| % Male | 100 | 69 |

*Note: one pathologist (W. A. Wells) is excluded from this table (ineligible to participate in slide read, but contributes reports to the database).

Table 2. Distribution of diagnoses (n) by slide for the 30 representative cases

| Slide | Benign (n) | Benign with atypia (n) | Noninvasive malignant (n) | Invasive Malignant (n) |
|---|---|---|---|---|
| A | 26 | 0 | 0 | 0 |
| B | 26 | 0 | 0 | 0 |
| C | 26 | 0 | 0 | 0 |
| D | 26 | 0 | 0 | 0 |
| E | 26 | 0 | 0 | 0 |
| F | 26 | 0 | 0 | 0 |
| G | 24 | 0 | 0 | 0 |
| H | 25 | 1 | 0 | 0 |
| I | 25 | 1 | 0 | 0 |
| J | 25 | 1 | 0 | 0 |
| K | 25 | 1 | 0 | 0 |
| L | 25 | 1 | 0 | 0 |
| M | 25 | 1 | 0 | 0 |
| N | 24 | 1 | 0 | 1 |
| O | 23 | 1 | 1 | 0 |
| P | 23 | 1 | 0 | 1 |
| Q | 22 | 3 | 1 | 0 |
| R | 22 | 4 | 0 | 0 |
| S | 19 | 6 | 1 | 0 |
| T | 13 | 12 | 1 | 0 |
| U | 13 | 12 | 1 | 0 |
| V | 0 | 1 | 25 | 0 |
| W | 0 | 0 | 1 | 25 |
| X | 0 | 0 | 6 | 20 |
| Y | 0 | 0 | 13 | 12 |
| Z | 0 | 0 | 16 | 10 |
| AA | 0 | 0 | 0 | 26 |
| BB | 0 | 0 | 0 | 26 |
| CC | 0 | 0 | 0 | 26 |
| DD | 0 | 0 | 0 | 26 |

diagnostic categories, but particularly for distinction between the benign and malignant categories, between the invasive malignant category and all other categories, and between the benign (without atypia) category and all other categories. This is a higher level of agreement than was reported in a prior study of diagnostic reproducibility of proliferative breast lesions *(4)*. The slides reviewed in that study *(4)* were selected to include a high proportion of controversial and difficult borderline lesions; our slides comprised a representative sample of the diagnostic categories seen routinely in a general pathology practice. The participants in the prior study also used mutually agreed on diagnostic criteria while our participants followed their individual criteria for diagnosis within a standardized checklist.

Despite the excellent agreement overall, there are situations when anything less than perfect agreement may be clinically unacceptable. A diagnosis of cancer, when none is present, may result in unnecessary therapy and concern. Similarly, misdiagnosing cancer as a benign condition would result in needed therapy not being received. In this study, such critical disagreements occurred primarily in the differentiation between diagnoses of benign with atypia and noninvasive malignant. In most institutions, a woman whose breast biopsy diagnosis is benign with atypia receives follow-up surveillance and no treatment, whereas a noninvasive malignant diagnosis warrants at least surgical excision and often more extensive treatment *(2)*. Among the 30 reviewed cases in our study, five (8%) of 66 diagnoses of noninvasive malignant (cases O, Q, S, T, and U) represent instances where the consensus opinion of the other pathologists was that no cancer was present. In seven instances of a noninvasive malignant diagnosis (cases W and X), most pathologists had diagnosed invasive cancer; in two cases (Y and Z), pathologists were approximately equally divided between invasive and noninvasive assessments. There were two instances of a diagnosis of invasive malignant for which the consensus opinion was no cancer (cases N and P), and one instance of a diagnosis of no cancer (benign with atypia, case V) where the consensus opinion was that cancer (noninvasive) was present. Most pathologists in our state have told us they confer with their colleagues in difficult diagnostic breast cases; therefore, these disagreements, usually representing the divergent view of one pathologist, would almost certainly have been exposed by a second evaluation. Disagreements might also be reduced through use of standardized diagnostic criteria for the differentiation between benign with atypia and noninvasive malignant categories *(4)*. Since only 30% of the pathologists in New Hampshire evaluate image-guided core biopsy specimens, the exceptional diagnostic agreement for these specimens throughout the state suggests that fears of a prolonged learning curve for the evaluation of such biopsies by pathologists when a stereotactic or ultrasound-guided service is introduced are unfounded.

Our study is one of few that have focused on the diagnostic reproducibility of routinely practicing pathologists without a special interest or expertise in diagnostic breast pathology. The most comprehensive study evaluating consistency of histopathologic reporting was carried out

Table 3. Kappa coefficients* for randomly selected slides in the four diagostic categories

| Diagnostic category comparisons | All slides (n = 30) | Image-guided core biopsy specimen slides (n = 9) | Excisional or mastectomy specimen slides (n = 21) |
|---|---|---|---|
| Benign versus malignant† | 0.95 | 0.98 | 0.94 |
| Benign without atypia versus all other categories | 0.79 | 0.94 | 0.73 |
| Benign with atypia versus all other categories | 0.22 | —‡ | 0.21 |
| Noninvasive malignant versus all other categories | 0.59 | 0.57 | 0.60 |
| Invasive malignant versus all other categories | 0.85 | 0.83 | 0.85 |

*There were 24 to 26 independent reviews per slide.

†$P<.001$ for all kappas unless otherwise noted.

‡Note that none of the nine slides had final diagnoses of benign with atypia.

by the United Kingdom National Breast Screening Programme in 1994 and involved up to 251 pathologists reviewing multiple sets of slides over 3 years *(5)*. As in our study, a high level of diagnostic consistency was achieved for most major categories of breast disease except when distinguishing benign with atypia and noninvasive, malignant categories. However, the slide sets did not represent the routine breast pathology caseload and slide quality was not formally assessed. The study of Bianchi et al. *(6)* showed good overall diagnostic agreement among 12 community-based Italian pathologists with comparable diagnostic discrepancies between benign with atypia and noninvasive malignant. However, although the study did control for the technical quality of the histologic sections, the cases selected for review were known to present diagnostic problems rather than randomly selected cases. In 1985, similar conclusions regarding diagnostic consistency were drawn from the study by members of the Medical Research Council Breast Tumor Pathology Panel in the U.K. who evaluated 40 consecutive cases submitted from health districts throughout the U.K. *(7)*.

Until more specific differentiating morphometric criteria or a biologic marker are determined, borderline proliferative breast lesions (representing 10% of our pathology database) will continue to be interpreted variably by community-based and expert pathologists alike. The natural history of low-grade noninvasive lesions as compared with the benign but atypical lesions is poorly understood. If the outcome of future clinical trials is to recommend comparable treatments for these borderline lesions, then the necessity to distinguish reproducibly between them may be alleviated.

Large cooperative clinical trials, such as the National Surgical Adjuvant Breast and Bowel Project, have tried to minimize inconsistencies of their pathologic findings by requiring that a central laboratory review all pathologic materials submitted by institutional pathologists *(9)*. Unless the clinical trials are specifically focusing on known areas of diagnostic variation, this procedure may not be necessary if the results of our current New Hampshire study apply broadly to pathologists elsewhere.

Two studies *(10,11)* have stated that optimal tissue fixation and processing are major factors in improving interobserver agreement in the histologic grading of breast carcinomas. In our study, reduced slide quality did not appear to affect diagnostic accuracy; indeed, for slides classified as of unsatisfactory interpretive quality or rated by greater than or equal to 25% as only satisfactory, the kappa coefficient improved from 0.64 to 0.69.

Three potential limitations of this study merit consideration. First, while the participation rate was good (80% of eligible pathologists submitting information to the pathology database and completing some aspects of the study), only 59% completed the slide review portion of the study. Willingness to take part in such a slide review may be considered a potential bias in participant selection and result in increased accuracy and agreement as compared with the community as a whole. Second, just one representative slide per case was requested for review, increasing the potential for sampling variability. In routine daily practice, pathologists would evaluate more than one slide from excisional and mastectomy specimens. Third, the uniform reporting form may have influenced final interpretations, since its format discouraged wordy comments.

In summary, breast pathology diagnoses among community pathologists in New Hampshire are highly reliable overall, particularly for the benign versus malignant categories, and for core biopsy specimens and special type invasive tumors. Tissue processing and slide quality do not measurably affect diagnostic agreement. Rereading breast pathology cases in categories critically important for determining treatment plans (benign with atypia and noninvasive malignant categories) only occurs in about 74% and 79% of the cases, respectively. A consistent slide review policy for breast pathology could lessen the likelihood of misclassification error. Clinically relevant diagnostic disagreements still occur, however, among noninvasive malignant diagnoses. The willingness of so many New Hampshire pathologists to participate in this project attests to their continued commitment to address these diagnostic variations and minimize clinically significant disagreements.

## References


(1) Ernster V, Barclay J, Kerlikowske K, Grady D, Henderson C. Incidence of and treatment for ductal carcinoma in situ of the breast. JAMA 1996;12:913–8.

(2) Choi WS, Parker BA, Pierce JP, Greenberg ER. Regional differences in the incidence and treatment of carcinoma in situ of the breast. Cancer Epidemiol Biomarkers Prev 1996;4: 317–20.

(3) Rosai J. Borderline epithelial lesions of breast. Am J Surg Pathol 1991;15:209–21.

(4) Schnitt SJ, Connolly JL, Tavassoli FA, Fechner RE, Kempson R, Gelman R, et al. Interobserver reproducibility in the diagnosis of ductal proliferative breast lesions using standardized criteria. Am J Surg Pathol 1992;16:1133–43.

(5) Sloane JP, Ellman R, Anderson TJ, Brown CL, Coyne J, Dallimore NS, et al. Consistency of histopathological reporting of breast lesions detected by screening: findings of the U.K. National External Quality Assessment (EQA) Scheme. U.K. National Coordinating Group for Breast Screening Pathology. Eur J Can 1994;30A:1414–9

(6) Bianchi S, Palli D, Galli M, Arisio B, Cappa A, Dal Forno S, et al. Reproducibility of histological diagnoses and diagnostic accuracy of nonpalpable breast lesions. Pathol Res Pract 1994; 1:69–76.

(7) Beck JS. Observer variability in reporting of breast lesions. J Clin Pathol 1985;38:1358–65.

(8) Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometries 1977;33:159–74.

(9) Fisher ER, Costantino J. Quality assurance of pathology in clinical trials. The National Surgical Adjuvant Breast and Bowel Project experience. Cancer 1994;74(9 Suppl):2638–41.

(10) Robbins P, Pinder S, de Klerk N, Dawkins H, Harvey J, Sterrett G, et al. Histological grading of breast carcinomas: a study of interobserver agreement. Hum Pathol 1995;26:873–9.

(11) Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 1991;19:304–10.


## Notes


Supported by The Centers for Disease Control to the New Hampshire Division of Health and Human Services (411216); by the U.S. Department of Defense (DASD17-94-J-4109); and by Public Health Service grant CA23108 from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services.

We thank Dr. David Donsker for encouraging this project and Julie Wade, Dr. Marguerite M. Stevens, and Sarah Meyers for editorial assistance. We also greatly appreciate all New Hampshire pathologists who took part in this Project.

Manuscript received June 12, 1997; revised October 15, 1997; accepted October 29, 1997.

Journal of Community Health, Vol. 23, No. 2, April 1998

# CAN WE IMPROVE BREAST PATHOLOGY REPORTING PRACTICES? A COMMUNITY-BASED BREAST PATHOLOGY QUALITY IMPROVEMENT PROGRAM IN NEW HAMPSHIRE

Patricia A. Carney, PhD; M. Scottie Eliassen, MS; Wendy A. Wells, MD; William G. Swartz, MS

ABSTRACT: We implemented a regional quality assurance program in New Hampshire (NH) to evaluate breast pathology practices and attempt to improve the completeness of information provided in breast surgical pathology reports. We also assessed the degree to which NH pathologists agree with National Guidelines. The program's objective was to promote a consistent standard of care for patients whose breast pathology is interpreted in NH. Using a sequential survey technique, we were able to obtain consensus on breast tissue report content that was similar to National Guidelines. We also found that 52% of the reporting elements improved in the post-intervention period, although only one reached statistical significance. In conclusion, pathology interpretation is the "gold standard" for determining both screening effectiveness and subsequent treatment of breast cancer, yet variability in breast tissue reporting exists. It is critical that more research be done to improve breast pathology interpretation and reporting practices.

## INTRODUCTION

Research in breast cancer screening and diagnosis has received a great deal of recent attention as the effectiveness of screening mammography in women of various age groups is questioned.[1-5] New Hampshire (NH) is one of ten states currently in the process of developing a popula-

Patricia A. Carney is Research Assistant Professor of Community and Family Medicine, M. Scottie Eliassen and William G. Swartz are Research Associates in the Dept. of Community and Family Medicine, and Wendy A. Wells is Assistant Professor of Pathology, all at Dartmouth Medical School and the Norris Cotton Cancer Center, Hanover and Lebanon, NH.

Requests for reprints should be addressed to Patricia A. Carney, PhD, Norris Cotton Cancer Center, Dartmouth Medical School, 1 Medical Center HB 7925, Lebanon, New Hampshire 03756.

This Project was funded by the New Hampshire Division of Public Health Services through a cooperative agreement with the Centers for Disease Control grant # U57-CCU108362-02 and by the United States Department of Defense (DAMD17-94-J-4109). The authors gratefully acknowledge the enthusiastic support of the participating New Hampshire pathologists and the N.H. Society of Pathologists. We also extend our sincerest appreciation to Julie Wade for editorial assistance; Brenda Berube for her administrative expertise in assisting with this study; and Margaret Murphy at the New Hampshire Division of Public Health Services for her assistance in obtaining funding.

tion-based mammography registry (New Hampshire Mammography Network).[4] NH is also a state with a large Centers for Disease Control funded community-based breast and cervical screening program that is supplemented by state funds. In combination, these programs will provide 4,000 free mammograms to underserved women. Such screening programs are proliferating in virtually all states around the country.

Because the pathological diagnosis of a breast lesion is traditionally considered the "gold standard" in evaluating screening effectiveness and determining treatment modalities, follow-up for the registry tracking system and the State screening programs includes obtaining pathology reports on all breast tissue examined and linking these to mammographic interpretations. To evaluate the completeness of breast surgical pathology reports and diagnostic accuracy, we implemented a regional breast pathology quality improvement (QI) program in NH. The objective of the program was to promote a consistent reporting standard and improve breast tissue reporting for patients whose breast pathology is interpreted and reported within the state.

The QI program had two phases. In Phase I we conducted a baseline assessment of current practices in specimen sources, specimen evaluation, slide preparation and pathology reporting in NH hospitals. We additionally established state-wide consensus of diagnostic core variables for breast pathology reports based on nationally established criteria[5,6] and assessed whether the process of the pathologists' coming to consensus improved subsequent report content. In Phase II we determined the degree of agreement amongst pathologists in the diagnostic assessment of breast tissue. We also explored the degree to which variability in diagnostic interpretation is associated with sample sources, specimen evaluation or slide preparation. The results of Phase II are reported in detail elsewhere.[7] This paper describes the activities undertaken in Phase I.

## METHODS

### Physician Recruitment, Survey Development and Implementation

Pathologist eligibility requirements included interpreting breast tissue pathology in a NH practice and not relocating or retiring within the study time period (one year). Because the QI program contained an extensive evaluation component, Institutional Review Board approval was applied for and granted. The QI Study was described in detail in subsequent letters and fact sheets, and informed consent was obtained from all pa-

thologists willing to participate. In addition, the study's pathology liaison (WAW) visited each pathology lab in the state to discuss the program's objectives personally.

Three surveys were then designed, developed and implemented. One obtained information on the demographic and practice characteristics of pathologists, which was administered after participants' informed consents were received by the Project office. The second survey ascertained specimen sources and methods of preparation and processing by participating laboratories. This was administered to one designated pathologist at each laboratory. The final survey ascertained which diagnostic criteria pathologists felt should routinely appear in a breast pathology report.

The surveying of report content began after pathology report baseline data collection was complete (see below). A sequential surveying technique was utilized to obtain state-wide overall agreement on the content of such reports:

- the initial survey was administered, asking pathologists what components they felt should routinely appear in a breast pathology report, according to sample source and diagnosis;
- data from all surveys were entered and analyzed using descriptive statistics;
- a draft of the results was sent to participating pathologists with a request for feedback;
- pathologists' comments were compiled and the checklist revised;
- the revised checklist was mailed to participating pathologists with another request for their comments;
- when pathologists' comments were no longer substantive, the checklist was finalized and circulated for final approval;
- the final checklist was printed on pocket-size cards and distributed to all pathologists in the state.

### Pathology Report Database Design, Data Entry, and Quality Assurance

As part of the NHMN mammography registry project, the majority of women who obtain mammograms (approximately 90%) at participating facilities (n = 36) have agreed to allow access to their breast tissue reports. Institutional Review Board approval was obtained to maintain an anonymous database of breast pathology reports for women who did not consent to take part in the NHMN Project or who received mammograms at facilities not yet taking part in the Project.

At each institution participating in the Breast Pathology QI Project, a designated pathologist or laboratory assistant made copies of all breast tissue reports (including fine needle aspirates) and submitted them, in batched quantities, to the project office. Breast tissue reports were initially collected for a three month period to assess baseline content of breast pathology reports. These were abstracted by MSE and entered into a specially designed relational database.

The database was developed by the study's pathology liaison (WAW) and pathology coordinator (MSE), using the core variables designated by the National Cancer Institute Sponsored Breast Cancer Surveillance Consortium[8] and other information commonly included in pathology reports in New Hampshire.[7] To maintain confidentiality, no identifying information was included in the database. Each patient, pathologist, and lab was assigned a unique ID used for linking and tracking data.

Data collected in the pathology database included: data links (anonymous and unique patient ID, patient's date of birth and gender); site information (lab code, pathologist code); case information (date of procedure, case number, type of procedure and laterality, history of previous biopsies); and diagnostic information (includes a number of categories for both benign and malignant conditions, as well as prognostic indicators such as Scarff-Bloom-Richardson (SBR) grade and estrogen or progesterone status).

In the initial stages of database design and data collection, information from submitted pathology reports was transcribed onto a standard paper form and reviewed for accuracy by the pathology liaison (WAW) prior to entry into the pathology database. When the format of the database stabilized, a transition was made to entering data directly into the computer from the pathology reports. To evaluate the accuracy of information extraction from the reports and data entry, 20 records from every batch of 100 sequentially entered in the database were randomly selected for review by the pathology liaison (WAW).

Percent agreement between the two observers (MSE and WAW) on the randomly selected records entered to date ($n=160$) is between 75 and 100% with a mean of 91%. The inconsistencies between the reviewers were minor in every case. Two discrepancies led to further refinement of the database to accommodate additional diagnostic criteria commonly reported in the state. The remaining errors were as follows: lesion size not recorded ($n=5$); histological subtype recorded incorrectly ($n=3$); benign microcalcifications excluded ($n=2$); intraductal papillomatosis recorded as single papilloma ($n=1$); lymph node counts did not tally ($n=1$); omit-

ted lobular hyperplasia (n=1); type of invasion recorded incorrectly (n=1).

After the baseline period was complete (study months 1–6) and sequential administration of the report content survey had begun, pathology reports continued to be batched and sent by participating labs throughout the study time period. A continuing medical education session was held in the ninth study month to share the results of the data collected to date, particularly the results of state-wide consensus on breast pathology report content. Results of interpretive agreement from Phase I were also shared .[7]

### Assessing Improvements in Breast Pathology Reports

To assess whether breast pathology report content improved as a result of coming to consensus on content, we randomly selected 45 reports of invasive and non-invasive breast cancer based on their relative distributions in the database in the baseline period and compared them to 45 reports of comparable distribution (invasive/non-invasive) randomly drawn from the database after the sequential surveying technique was implemented. Comparisons were made based on a reporting variable being mentioned as either present or absent in the report versus no mention of relevant variables (either as present or absent) in the baseline versus post survey periods. Descriptive statistics and the McNemer's test of symmetry were used to evaluate improvements in report content.

## RESULTS

### Characteristics of Pathologists and Laboratories

The demographic/practice characteristics survey and the report content survey were completed by 91% and 94% of participating pathologists, respectively. The survey on specimen preparation was completed by 83% of designated pathologists, representing the 14 participating labs where breast tissue is processed.

Forty-three pathologists interpret breast pathology in New Hampshire and were eligible to take part in the Project. Of these, 35 (79%) agreed to participate. Seventeen of the state's 26 hospitals have laboratories where breast specimens are grossed in and read; 14 (82%) agreed to take part. Ten hospitals have labs that cut slides; 8 (80%) took part.

Project participants ranged in age from 31 to 60 with a mean age

of 47 (S.D. = 8.0 years). The majority were male (72%). The mean year of graduation from medical school was 1976 with a range between 1958 and 1989. The mean year for completion of residency programs was 1981 with a range between 1963 and 1994. Thirty six percent of participating pathologists underwent fellowship training and completed this training between 1982 and 1995. Ninety-seven percent were Board certified in pathology. Pathologists had been practicing at their current laboratory locations for between 3 months and 33 years with a mean of nine years (S.D. = 8.2 years). Pathologists had been interpreting breast pathology for 2–37 years with a mean of 14 years (S.D. = 8.7 years). Lastly, they participated in 15–191 hours of continuing medical education in pathology over the past year, with a mean of 76 hours (S.D. = 46 hours); this broad range is due to the mix of academic and community pathologists in the state.

The fourteen pathology laboratories reported reading between 700 and 17,280 pathology cases per year (mean = 5,241, S.D. = 3,820). Of these, between 20 and 720 cases per year are breast tissue (mean = 258, S.D. = 183). Ninety-three percent of sites evaluate fine needle aspirations at an annual volume of between 10 and 224 cases (mean = 74, S.D. = 63), and 29% reported evaluating stereotactic-guided core biopsies at an annual volume of between 5 and 104 cases (mean = 70, S.D. = 46).

At 64% of the labs, breast biopsies resulting from clinically detected masses or abnormal mammograms were always received in the fresh state from the operating room. In the remaining cases they were sometimes received fixed in formalin. A frozen section was performed on between 3 and 50% (mean 20% S.D. = 16%) of labs' breast biopsies. In 50% of labs, mammographic x-rays always accompanied excisional and/or needle localization specimens from the operating room, and 93% of pathologists found these accompanying films useful. In 86% of laboratories, specimen radiography was performed, and of these 8% were done in pathology and 92% were done in radiology.

At 93% of pathology labs in New Hampshire, excisional and/or needle localization specimens were always inked. For 71% of labs, fresh tissue (if present in adequate quantities) was submitted for biochemical assays for estrogen receptor and progesterone receptor status in all cases of malignancy; all of these sites use out-of-state labs for ER/PR. If diagnostic tissue was found to be limited, immunohistochemical studies for estrogen and progesterone receptivity were performed on paraffin-embedded blocks by all labs in all cases of malignancy. Twenty-one percent performed the immunohistochemical assays on-site; the remainder were sent to commercial labs. Forty-three percent of labs performed cell cycle analysis by flow cytometry in all cases of malignancy. Of these, 21% performed this

on-site with 36% performing this on fresh tissue and 57% performing it on paraffin-embedded tissue blocks.

### Opinions About Breast Tissue Report Content

All pathologists agreed that the presence of microcalcifications and epithelial hyperplasia (with and without atypia) should be mentioned in breast reports for benign disease. Ninety-three percent felt that biopsy size should be included, but few felt that information in the report regarding risk for development of subsequent cancer or follow-up recommendations was required (35% and 24% respectively).

Table 1 outlines the proportion of NH pathologists who advocate certain core diagnostic variables in breast pathology reports for non-invasive and invasive carcinoma; these are compared to the recommendations of the Association of Directors of Anatomic and Surgical Pathology (ADASP).[5,6] Here the range of recommended core diagnostic variables is 10–100% with biopsy and lesion size, whether it was discrete or multifocal, the in-situ pattern, presence of microcalcifications, margin status, and nipple involvement being advocated by more than 90% of pathologists for non-invasive carcinomas. Recommendations regarding prognostic risk or follow-up are advocated by only 14% of pathologists. Similar findings are noted for reporting on invasive carcinoma, though tumor histological type, tumor grade, and presence of associated extensive in-situ pattern, angiolymphatic and perineural invasion, and axillary lymph node dissections are additionally advocated by 100% of NH pathologists.

### Actual Performance on Content of Breast Tissue Reports

Table 2 illustrates our pre-post assessment of breast tissue reporting for invasive and non-invasive breast carcinoma. The variables in this table represent the core diagnostic variables participating NH pathologists agreed upon as part of the survey sequencing process. Here the range of core diagnostic variables reported in the baseline period range from 0–100, with size of excised specimen and laterality of the breast being the only core variables actually being reported on in more than 90% of the reports selected. The range is the same in the post sequencing survey period. Type of procedure done and resection margin status were reported in 89% of the reports in the post survey period.

Table 2 also indicates that more than half (52%) of the core diagnostic variables evaluated improved in the post survey period compared to baseline (those bolded in Table 2). However, only reporting on the extent of associated in-situ component was found to be statistically significant

TABLE 1

Percent of NH Pathologists Who Feel These Core Diagnostic Variables Should Be Routinely Included in All Breast Pathology Reports for Non-Invasive and Invasive Carcinoma, and ADASP Recommendations

| *Variables* | *% Say Report in Non-Invasive Carcinoma* | *% Say Report Invasive Carcinoma* | *ADASP Recommends* |
|---|---|---|---|
| GROSS DESCRIPTION: | | | |
| Biopsy size | 100 | 100 | yes |
| MICROSCOPIC DESCRIPTION: | | | |
| Lesion size | 90 | 93 | yes |
| Maximum diameter (cm) | 83 | 76 | at least |
| Two dimensions (. . . . x. . . . cm) | 35 | 41 | * |
| Three dimensions (. . . . x. . . . x. . . . cm) | 55 | 62 | preferred |
| Tumor histological subtype | — | 100 | yes |
| Tumor grade (e.g.: Scarff-Bloom-Richardson) | — | 100 | yes |
| Discrete or multifocal | 100 | 100 | * |
| Presence of associated extensive in-situ component | — | 100 | yes |
| Estimation of % of the total tumor size | — | 76 | * |
| In-situ pattern | 100 | — | yes |
| Presence of Microcalcifications | 97 | — | mammo correlation |
| Benign association | 52 | 69 | |

TABLE 1 (*Continued*)

| | | | |
|---|---|---|---|
| Malignant association | 62 | 72 | |
| Reserction Margin (RM) status | 100 | — | yes |
| Involvement by infiltrating carcinoma | — | 100 | yes |
| Involvement by in-situ carcinoma | — | 97 | yes |
| Distance between tumor and closest RM | 76 | — | yes |
| Involvement of dermal lymphatics | — | 93 | * |
| Axillary LN dissections (positive vs. negative) | — | 100 | yes |
| Angiolymphatic and perineural invasion | — | 100 | yes (perineural optional) |
| Involvement or not of nipple (Paget's) | 93 | — | yes |
| Correlation with previous biopsies | — | 93 | * |
| ER/PR status | 72 | — | optional |
| Biochemical assay | 62 | 72 | optional |
| Immunohistochemical evaluation | 83 | 100 | optional |
| Flow cytometric cell cycle analysis | 35 | 45 | optional |
| TNM classification | — | 69 | optional |
| Specification of different components of FCD | 76 | 69 | ADH, papillomas |
| Presence of a mononuclear cell infiltrate | — | 31 | * |
| Presence of necrosis | — | 83 | * |
| Recommendations regarding prognostic risk | 14 | — | * |
| Recommendations regarding follow-up | 14 | 10 | * |

*Not addressed by the ADASP.
— Not relevant for that category or not asked on subsequent surveys.

TABLE 2

Assessment of Breast Tissue Reporting for Invasive and Non-Invasive Breast Carcinoma at Baseline and Post Sequencing Survey

| *Variables* | *% at Baseline* | *% at Post Sequencing Survey* | *p value* |
|---|---|---|---|
| GROSS DESCRIPTION: | n=45 | n=45 | — |
| • All resection margins inked | 56 | 41 | 0.18 |
| • Biopsy size | 93 | 100 | — |
| • Laterality of breast | 100 | 100 | — |
| Procedure done | 80 | 89 | 0.29 |
| MICROSCOPIC DESCRIPTION: | | | |
| • Tumor size: Max. diameter | 72 | 78 | 0.60 |
| • Tumor grade (e.g. Scarff-Bloom-Richardson) | 79 | 79 | 1.00 |
| • Associated in-situ component: | 73 | 73 | 1.00 |
| a) Extensive/Not extensive | 50 | 88 | 0.01* |
| b) Pattern(s) | 4 | 0 | — |
| • Microcalcifications Benign/Malignant association | 22 | 42 | 0.60 |
| • Resection Margin (RM) status | 78 | 89 | 0.25 |
| Involvement by invasive/non-invasive Ca | 16 | 16 | 1.00 |
| Distance from closest RM (not for lobular Ca) | 71 | 42 | 0.16 |
| • ER/PR status: Immunohistochemical/Biochemical | 47 | 36 | 0.32 |
| To be mentioned, if present: | | | |
| Axillary Lymph Nodes (positive Vs negative) | 27 | 36 | 0.48 |
| Angiolymphatic (incl. dermal) and perineural invasion | 54 | 66 | 0.71 |
| Involvement of nipple (Paget's) | 60 | 80 | 0.56 |
| Correlation with previous biopsies/cytology specimens | 31 | 38 | 0.53 |
| NON-INVASIVE ONLY: | | | |
| • In-situ pattern(s) | 40 | 29 | 0.89 |
| • Discrete or multifocal | 0 | 2 | — |

TABLE 2 (*Continued*)

| | | | |
|---|---|---|---|
| • Nuclear Grade | 0 | 0 | — |
| • No Invasion Seen | 2 | 4 | — |
| OPTIONAL INCLUSIONS: | | | |
| Flow cytometric cell cycle analysis | 11 | 8 | 0.71 |
| • TNM classification | 24 | 13 | 0.29 |
| • Specification of different FCD components | 32 | 35 | 0.82 |

*$p < 0.05$

($p \leq 0.01$). Four report elements remained unchanged, and six were actually reported less often in the post survey period than they had been at baseline.

## DISCUSSION

We observed high levels of interest in our breast pathology QI project by NH pathologists and laboratories, as indicated by our high response rates (79% and 82% respectively). Clearly this is an important issue for pathologists in the state. Our study revealed that NH pathologists are well trained and experienced, all completing a residency training program and nearly all being Board certified. In the last 15 years, 35% of the pathologists had acquired additional Fellowship training. As well as evaluating routine surgical excisional biopsies, including needle localization specimens, diagnoses were made on stereotactic- and ultrasound-guided core biopsies and fine needle aspirations.

We also learned that a great deal of variability exists in the volumes of breast pathology interpreted in NH laboratories. Only one participating laboratory was based in an academic medical center; the others were small to medium sized community-based hospitals in a mix of urban and rural areas. We found essentially no commercial laboratories are used to process breast tissue (hospitals in one region of the state use an independent local laboratory), except to determine estrogen/progesterone status and to perform immunohistochemical assays.

The procedures undertaken to process specimens vary somewhat. The concentration of formalin used for tissue fixation, the time of fixation, sectioning thickness, and tissue staining characteristics are the most variable criteria amongst different laboratories and if substandard, can cause interpretive variations in diagnosis. However, the results of Phase II of this

project indicate that within NH, there was no appreciable variability in diagnostic interpretation associated with sample sources or slide preparation.[7]

The College of American Pathologists (CAP) provides regular surveys (Q-probe studies) which are designed to measure service quality in individual laboratories as compared to the performance of other participating institutions across the country. The results of a recent Q-probe study (95–03) analyzed how many pathologists are already standardizing the processing of their specimens and the diagnostic and prognostic information detailed in their surgical pathology reports.[9] Four hundred and thirty-four pathology laboratories participated in the study nationwide. The variability was marked. Most participants (65.7%) admitted that they do not use standardized checklists to report core diagnostic variables. The handling of breast biopsy specimens was greatly influenced by how the tissue was received in the laboratory (fresh or fixed), the clinical information provided, and the presence or abscence of a radiograph. Overall, breast biopsy specimen handling in NH fell at about the 70th percentile relative to the performance of the other participating institutions in this Q-probe.

Recently, the Association of Directors of Anatomic and Surgical Pathology (ADASP) published recommendations regarding core variable diagnostic features that should be included in all surgical reports for breast carcinoma.[5] A standardized approach to the gross evaluation and tissue processing of breast excision specimens has also been detailed.[6,10] The recommendations were intended as an educational resource rather than a compulsory requirement, but it was hoped that the suggestions would lead to more standardized information being provided to clinicians for them to better evaluate prognostic predictors, disease staging and therapy.

Interestingly, the information that clinical physicians (radiation oncologists, surgeons, oncologists and radiologists) regularly sought on breast pathology reports was also evaluated in the 1995 CAP Q-probe (95–03). Between 76 and 95% of clinical physicians desired that diagnostic and prognostic criteria similar to those detailed by the ADASP be included routinely in all breast carcinoma pathology reports as necessary factors in evaluating optimal patient care. The main concluding recommendation of the study was that a checklist of diagnostic core variables, approved by both the pathologist and the involved physicians, should be included in breast pathology reports.

We were pleased to achieve consensus with participating pathologists on the core diagnostic variables that should be present when a breast

cancer (either invasive or non-invasive) is diagnosed and that overall NH pathologists are in agreement with ADASP guidelines. We were also pleased to show improvements in more than half the reporting elements under study; however, we were disappointed that statistical improvement was only noted in one of the reporting elements agreed upon.

Several areas warrent further study and discussion. First, the resources available to conduct the report content assessment were minimal. A total of 90 reports, 45 in the baseline period and 45 in the post sequencing survey period could only have provided enough power to detect a large effect size. A larger sample size may have identified statistical differences in report content between the two time periods. This is certainly an area for future study.

Second, we suspect that there are characteristics of pathology specimens that promote reporting the absence or presence of certain features, which may have affected our findings. We also suspect that it may be much easier for a pathologist to be prompted by the presence of a diagnostic variable during interpretation and reporting than it is to report the absence of that same variable, regardless of its significance. As part of our project, we developed laminated pocket-sized cards with the agreed upon core diagnostic variables listed. We hoped that the cards would assist in prompting the pathologists to be more consistent in their reporting; this appears not to be the case. Most NH pathologists very likely do not specialize in breast tissue interpretation and the process of using or not using these cues to action based on the variety of tissue being interpreted could effect the impact of such an intervention. Certainly, more research is needed to understand factors that influence breast tissue reporting.

We noted that providing information on the text report for prognostic risk and making follow-up recommendations was only advocated by between 10–14% of NH pathologists. Though we expect that many pathologists would agree that noting prognostic risk as well as follow up recommendations in their reports would be useful, these factors are likely best determined collaboratively by the pathologist, surgeon, radiologist and oncologist. Risk and recommendations are always discussed at length in settings such as the weekly tumor boards where subsequent treatment plans are discussed. We feel this may have influenced pathologists' not advocating these variables in their reports.

The pathologist's text report provides the basis for critical public health and cancer treatment decisions. More consistency is needed on breast tissue reporting than we were able to achieve in our study. This is an immense area for further study. Public health programs that study the ef-

fectiveness of mammography and/or that offer mammography screening services should implement quality assurance programs to monitor and attempt to reduce variability noted in pathology interpretation and reporting practices.

## REFERENCES


1. Hakama M, Pukkala E, Kallio M, Godenhjelm K and Svinhufvud U. Effectiveness of screening for breast cancer in women under 50 years at entry: the Kotka pilot project in Finland. *Int J Cancer* 1995; 63:55–7.
2. Lindfors KK and Rosenquist CJ. The cost-effectiveness of mammographic screening strategies. *JAMA* 1995; 274:881–4.
3. Fletcher SW, Black W, Harris R, Rimer BK and Shapiro S. Report of the International Workshop on Screening for Breast Cancer. *J Natl Cancer Inst* 1993; 85:1644–56.
4. Carney PA, Poplack SP, Wells WA and Littenberg B. The New Hampshire Mammography Network: The development and design of a population-based registry. *AJR Am J Roentgenol* 1996; 167:367–72.
5. Recommendations for the Reporting of Breast Carcinoma. Association of Directors of Anatomic and Surgical Pathology. *Am J Clin Pathol* 1995; 104:614–619.
6. Recommendations of the Association of Directors of Anatomic and Surgical Pathology: Part I. Immediate management of mammographically detected breast lesions. *Hum Pathol* 1993; 24:689–690.
7. Wells WA, Carney PA, Eliassen MS, Tosteson AN and Greenberg ER. A Statewide Study of Diagnostic Agreement in Breast Pathology. *J Natl Cancer Inst* 1998; 90:142–145.
8. Ballard-Barbash R, Taplin SH, Yankaskas BC, *et al.* Breast Cancer Surveillance Consortium: A national mammography screening outcomes database. *Am J Roentgenol, In Press*, 10/1997.
9. Nakhleh RE, Jones B and Zarbo RJ. Mammographically directed breast biopsies: a College of American Pathologists Q-Probes study of clinical physician expectations and of specimen handling and reporting characteristics in 434 institutions. *Arch Pathol Lab Med* 1997; 121:11–8.
10. Schnitt SJ and Connolly JL. Processing and Evaluation of Breast Excision Specimens: A Clinically Oriented Approach. *Am J Clin Pathol* 1992; 98:125–137.

The American Journal of Surgical Pathology 24(5): 651–659, 2000



# Pathologists' Agreement With Experts and Reproducibility of Breast Ductal Carcinoma-in-Situ Classification Schemes

Wendy A. Wells, M.D., Patricia A. Carney, Ph.D., M. Scottie Eliassen, M.S., Margaret R. Grove, M.S., and Anna N. A. Tosteson, Sc.D.

Several histologic classifications for breast ductal carcinoma in situ (DCIS) have been proposed. This study assessed the diagnostic agreement and reproducibility of three DCIS classifications (Holland [HL], modified Lagios [LA], and Van Nuys [VN]) by comparing the interpretations of pathologists without expertise in breast pathology with those of three breast pathology experts, each a proponent of one classification. Seven nonexpert pathologists in New Hampshire and three experts evaluated 40 slides of DCIS according to the three classifications. Twenty slides were reinterpreted by each nonexpert pathologist. Diagnostic accuracy (nonexperts compared with experts) and reproducibility were evaluated using inter- and intrarater techniques (kappa statistic). Final DCIS grade and nuclear grade were reported most accurately among nonexpert pathologists using HL (kappa = 0.53 and 0.49, respectively) compared with LA and VN (kappa = 0.29 and 0.35, respectively, for both classifications). An intermediate DCIS grade was assessed most accurately using HL and LA, and a high grade (group 3) was assessed most accurately using VN. Diagnostic reproducibility was highest using HL (kappa = 0.49). The VN interpretation of necrosis (present or absent) was reported more accurately than the LA criteria (extensive, focal, or absent; kappa = 0.59 and 0.45, respectively), but reproducibility of each was comparable (kappa = 0.48 and 0.46, respectively). Intrarater agreement was high overall. Comparing all three classifications, final DCIS grade was reported best using HL. Nuclear grade (cytodifferentiation) using HL and the presence or absence of necrosis were the criteria diagnosed most accurately and reproducibly. Establishing one internationally approved set of interpretive definitions, with acceptable accuracy and reproducibility among both pathologists with and without expertise in breast pathology interpretation, will assist researchers in evaluating treatment effectiveness and characterizing the natural history of DCIS breast lesions.

From the Department of Pathology (W.A.W.), Center for the Evaluative Clinical Sciences (M.R.G., A.N.A.T.), Department of Community and Family Medicine (P.A.C., M.S.E., M.R.G.), and Department of Medicine (A.N.A.T.), Dartmouth Medical School, Hanover, New Hampshire, U.S.A.

Address correspondence and reprint requests to Wendy A. Wells, MD. Department of Pathology, Dartmouth-Hitchcock Medical Center, 1 Medical Center Drive, HB 7999, Lebanon, NH 03756, U.S.A.; e-mail: Wendy.A.Wells@dartmouth.edu

Key Words: Breast—Ductal carcinoma in situ—Diagnostic reproducibility—Diagnostic accuracy.

*Am J Surg Pathol 24(5): 651–659, 2000.*

Ductal carcinoma in situ (DCIS) of the breast is a heterogeneous lesion with variable biologic behavior, which currently accounts for more than 20% of mammographically detected breast lesions.[9,14] In combination with tumor size and margin status,[33] mammographic correlation,[17] and other selected biologic markers,[3,5,6,19,25,28,39] histologic appearances help predict the clinical behavior of DCIS lesions. Because variations in clinical behavior most likely reflect the histologic heterogeneity of DCIS lesions, well-defined and reproducible criteria for the range of appearances of DCIS are necessary if the appropriate treatment is to be administered. Despite this, few of the published classifications (original or modified) attempt, at the time of publication, to assess the reliability/diagnostic reproducibility of the described criteria among pathologists with and without expertise in breast pathology interpretation.[32]

The objective of our study was to evaluate nonexpert pathologists' diagnostic accuracy and reproducibility of three DCIS classifications using inter- and intrarater techniques. We tested the overall diagnostic accuracy of the community pathologists by comparing their interpretations with those of three breast pathology experts (considered the diagnostic "gold standard"), each of whom is a proponent of one of the classification systems under study. We also evaluated the diagnostic reproducibility among nonexpert pathologists for each of the three classifications overall as well as the reproducibility of separate criteria within each classification.

## METHODS

After approval by the institutional Committee for the Protection of Human Subjects, three histologic classifi-

cations for DCIS were selected for evaluation: the Holland classification (HL),[18] the modified Lagios classification (LA),[32] and the Van Nuys scheme (VN).[33] A data collection instrument (standardized pathology reporting form) was designed to evaluate the prominent aspects of each DCIS classification system (Table 1) in a way that would allow for comparisons across all three classifications. In most cases these instruments were reviewed by the experts before implementation.

## Classification Criteria Evaluated

The HL, used by the European Pathologists Working Group, emphasizes cytonuclear differentiation primarily, and architectural differentiation (cellular polarization) secondarily. This system classifies DCIS into three groups: poorly, intermediately, and well differentiated. The term "comedo necrosis" is not used as a diagnostic criterion. Necrosis is defined as a frequently associated feature that may be present variably as central necrosis or as individual cell necrosis and autophagocytosis. Other frequently associated features are descriptive growth patterns and calcification type. The classification does not attempt to include the comparatively rare special types of DCIS, such as apocrine, mucinous, or signet ring types, because it is uncertain into which group these special types should be placed. In this study, participants formally assessed cytonuclear differentiation (nuclear grade) and cell polarization according to the published definitions. Final overall DCIS grade was also assessed. Other criteria (such as necrosis and growth patterns) were used additionally to arrive at the overall final differentiation but were not recorded formally.

The LA system classifies DCIS as high grade (high nuclear grade, extensive necrosis, comedo architecture), intermediate grade (intermediate nuclear grade, focal or absent necrosis, noncomedo architecture), or low grade (low nuclear grade, absent necrosis, noncomedo architecture). Special types of DCIS (pure apocrine and micropapillary types) are classified as a fourth option called "special-type." In this study, the participants formally assessed nuclear grade and necrosis (absent, focal, or extensive) according to the published definitions, and final overall DCIS grade. Growth patterns were not recorded formally.

The VN scheme evaluates the nuclear grade and the presence or absence of comedo-type necrosis. The presence of any high nuclear grade (with or without comedo-type necrosis) is defined as group 3. Of the remaining nonhigh-nuclear grade lesions, those with comedo-type necrosis are defined as group 2 and those without comedo-type necrosis are defined as group 1. Special types of DCIS are included in this classification. In this study, the participants formally assessed nuclear grade and presence or absence of "comedo-type" necrosis according to the published definitions, and final overall DCIS grade.

## Participating Pathologists

Three internationally recognized experts in breast pathology diagnosis (Michael Lagios, Rosemary Millis, and David Page), each a proponent of one of the three DCIS histologic classifications selected, agreed to provide the "gold standard" diagnosis according to their proposed classifications (Van Nuys [VN], Holland [HL], and Modified Lagios [LA], respectively). Seven of 44 eligible male and female pathologists (16%), representing differing geographic distributions, genders, and practice sizes in New Hampshire (NH), and without specific expertise in breast pathology, volunteered to participate in the study. Criteria for participation included at least 1 year of experience interpreting breast pathology in NH with no plans to retire or relocate within the study period.

## Study Cases

Using the New Hampshire Mammography Network's pathology database (described in detail elsewhere[8]), 50 patients with DCIS in the initial diagnosis were identified and obtained from the files of the study center (Department of Surgical Pathology, Dartmouth-Hitchcock Medical Center, NH) between 1992 and 1997.

TABLE 1. *Aspects of each DCIS classification system as recorded by the standardized reporting forms*

| System | Nuclear grade | Cell polarization | Necrosis | Final DCIS grade |
|---|---|---|---|---|
| Holland (HL) | • Poorly differentiated<br>• Intermediately differentiated<br>• Well differentiated | • Prominent<br>• Present, not prominent<br>• Absent/very focal | N/A | • Poorly differentiated<br>• Intermediately differentiated<br>• Well differentiated |
| Modified Lagios (LA) | • High<br>• Intermediate<br>• Low | N/A | • Extensive<br>• Focal<br>• Absent | • High grade<br>• Intermediate grade<br>• Low grade<br>• Special type |
| Van Nuys (VN) | • High<br>• Intermediate<br>• Low | N/A | • Present<br>• Absent | • Group 3<br>• Group 2<br>• Group 1 |

DCIS, ductal carcinoma in situ; N/A, not available.

From this set of 50 patients, 40 were selected randomly for this study. Baseline evaluation of the 40 slides used in the current study showed pure DCIS in 85% (n = 34) and DCIS in association with an invasive tumor in 15% (n = 6) of the study patients. Of the 40 patients evaluated in this study, only one patient was also used in our prior statewide study of agreement in NH of all diagnostic categories in breast pathology interpretation.[38] However, a different tissue block was used for this case of infiltrating carcinoma with extensive DCIS in each of the studies. Four tissue recuts, each 4 μm thick and stained with standard hematoxylin–eosin stain, were made of a representative block. Each recut was reviewed (by WAW) to ensure that the same histopathologic material was present. All patient and hospital or laboratory identifiers were removed, and a study code was applied to each slide. The four complete slide sets were mailed to each of the experts and the participating local pathologists according to a systematic rotation schedule.

During phase I of the study, the nonexpert NH pathologists were asked to evaluate each slide according to each of the three DCIS classifications, in a specific randomly assigned order. The original scientific papers detailing the three classifications, summaries of their diagnostic criteria, and the standardized pathology reporting form for the first classification in the specified order were also enclosed. To avoid rater fatigue and any bias introduced by interpreting criteria too closely together, participants were asked to assess all 40 slides for the first classification, then fax the completed pathology reporting form to the study coordinator (M.S.E.). On receipt of this fax, the appropriate reporting form for the next classification was sent. This was done until all slides were interpreted using all three classifications. To simulate usual working practices of the nonexpert pathologists, no teaching sets detailing the diagnostic criteria for each classification were distributed before the study. A set of study slides was sent to each of the three breast pathology experts, who evaluated each case according to their own proposed classification using the standardized pathology reporting forms (our data collection instrument also used by the NH pathologists).

Phase II of the project involved the seven NH pathologists reevaluating a set of 20 slides, selected randomly from the original set of 40. This was conducted 3 months after the completion of phase I. Again, each pathologist was asked to evaluate each slide according to the three classifications, using the same randomly assigned order as in phase I. Continuing Medical Education credits were awarded to all nonexpert pathologists completing the project, and each was sent a report comparing his or her individual interpretations with those of both the statewide aggregate and the experts.

### Statistical Analysis

Kappa statistics were used to assess agreement with experts, agreement between pathologists, and reproducibility within pathologists for each of the diagnostic categories within each classification system. The kappa sta-

**TABLE 2.** *Summary of agreement with the expert, and inter- and intrarater agreement among NH pathologists for criteria in each classification**

| | Agreement with expert (diagnostic accuracy) *k* (CI) | Agreement among NH pathologists (interrater agreement) *k* (CI) | Agreement within each NH pathologist (intrarater agreement) *k* (CI) |
|---|---|---|---|
| Holland (HL) | | | |
| Final DCIS differentiation | 0.53 (.28, .78) | 0.46 (.40, .51) | 0.49 (.19, .79) |
| Cytodifferentiation | 0.49 (.24, .73) | 0.45 (.39, .51) | 0.62 (.31, .91) |
| Cell polarization | 0.36 (–.03, .76) | 0.36 (.19, .53) | 0.43 (–.05, .90) |
| Modified Lagios (LA) | | | |
| Final DCIS grade | 0.29 (.06, .51) | 0.26 (.20, .31) | 0.57 (.29, .86) |
| Final nuclear grade | 0.35 (.11, .59) | 0.26 (.21, .32) | 0.67 (.38, .97) |
| Necrosis (extensive/focal/absent) | 0.45 (.21, .70) | 0.46 (.41, .52) | 0.63 (.33, .93) |
| Van Nuys (VN) | | | |
| Final DCIS group | 0.29 (.08, .50) | 0.26 (.20, .31) | 0.29 (.02, .56) |
| Final nuclear grade | 0.35 (.13, .58) | 0.29 (.23, .34) | 0.41 (.12, .70) |
| Necrosis (present/absent) | 0.59 (.30, .87) | 0.48 (.40, .55) | 0.67 (.33, 1) |

* Legend (24):

| Kappa statistic (*k*) | Interpretation of agreement |
|---|---|
| <0.4 | Poor to fair |
| 0.4–0.8 | Moderate to good |
| >0.8 | Excellent |

* Kappa statistics and 95% confidence intervals are shown.

TABLE 3. *Percent (no.) of cases classified into each DCIS final grade by the experts*

| | Poorly differentiated/ high grade/group 3 % (n) | Intermediately differentiated/ intermediate grade/group 2 % (n) | Well differentiated/ low grade/group 1 % (n) | Special type % (n) |
|---|---|---|---|---|
| Holland (HL) | 46 (17) | 32 (12) | 22 (8) | * |
| Modified Lagios (LA) | 40 (16) | 28 (11) | 30 (12) | 2 (1) |
| Van Nuys (VN) | 35 (14) | 28 (11) | 37 (15) | — |

DCIS, ductal carcinoma in situ.
* The Holland expert rated three slides as "special type"; however, they are omitted because the non-expert pathologists were not given an opportunity to rate them as such.

tistic evaluates level of agreement adjusted for agreement expected to occur by chance alone.[24] Kappa statistics less than 0.4 represent fair to poor agreement, values of 0.4 to 0.8 represent moderate to good agreement, and values of more than 0.8 represent excellent agreement (see legend for Table 2). To evaluate agreement with experts (that is, diagnostic accuracy), individual kappa statistics were estimated for each pathologist. A summary kappa value was obtained by combining the kappa values from individual pathologists using meta-analytic techniques.[11] Likewise, kappa values summarizing reproducibility of diagnostic classifications at repeat readings were estimated for each pathologist and combined across pathologists. To assess diagnostic agreement between pathologists, kappa statistics for multiple categories and multiple raters were estimated.[24] Statistical comparisons in agreement between classifications were made using a paired *t*-test of differences in kappa statistics for each pathologist. For each classification, we also estimated the proportion of slides for which the majority of community pathologists (four or more) agreed with the expert pathologist.[29] The comparisons between these proportions were made using a chi-square test.

## RESULTS

The seven NH pathologists ranged in age from 39 to 65 years old (mean, 49 ± 10 [standard deviation] yrs). They completed medical school 13 to 29 years before the study began (mean, 22 ± 11 yrs). Four of the seven participants (57%) completed fellowship training in addition to residency training; all seven participants are board certified. Six of the participants (86%) practice in community hospitals distributed throughout the state, and one pathologist practices in an academic medical center. None has a special interest in breast pathology. None of the participating pathologists (expert or nonexpert) disagreed with the overall diagnosis of DCIS in any of the 40 slides.

For the 40 patients we received a total of 840 diagnoses or three diagnoses per nonexpert pathologist per slide (3 × 7 × 40). The NH pathologists then provided an additional 418 of 420 review diagnoses when 20 of the cases were reinterpreted. The expert evaluating HL diagnosed special-type (apocrine) DCIS in three of the 40 patients. Because this option was not provided to the nonexpert pathologists (pure special types of DCIS are not designated as a separate final grade in the HL classification), then most of the missing data relates to the fact that the diagnoses of the nonexpert pathologists for these three cases could not be compared with the expert. Two of the three slides with missing diagnoses in the first round were also used during the second. In addition, two final diagnoses and one necrosis rating by nonexperts were missing in the second round.

Table 3 shows the percent of patients classified into each final DCIS grade by the experts according to each classification. Overall, the majority of patients showed high/poorly differentiated/group 3 final DCIS grade as defined by all three classifications, although the intermediate/group 2 and low/well differentiated/group 1 final DCIS grades were well represented in the study set. Compared with the three slides diagnosed by the expert using HL as special-type (apocrine) DCIS, a different slide was diagnosed as special-type (apocrine) DCIS by the expert using LA.

Table 2 summarizes agreement with the expert, and interrater (between pathologists) and intrarater (within pathologist) agreement among the nonexpert pathologists for the criteria evaluated in each classification. The diagnostic accuracy of nonexpert pathologists, compared with the experts, was considered moderate for the final DCIS grade and nuclear grade (cytodifferentiation) using HL (kappa = 0.53 and 0.49, respectively), and fair for final DCIS grade and final nuclear grade using LA and VN (kappa = 0.29 and 0.35, respectively, for both classifications). Final DCIS grade and final nuclear grade were best reproduced among the nonexpert pathologists using the HL criteria (kappa = 0.46 and 0.45, respectively).

Agreement with the expert for diagnosing cell polarization using HL was fair (kappa = 0.36). The diagnostic accuracy in distinguishing between either the presence or absence of necrosis using VN was moderate to good (kappa = 0.59) and slightly better than the three-tier system for evaluating necrosis (extensive, focal, or absent) according to LA (kappa = 0.45). Reproducibil-

ity among the NH pathologists for the two-tier (VN) and three-tier (LA) classifications of necrosis were comparable (kappa = 0.48 and 0.46, respectively). For all criteria in all classifications, the intrarater agreement was much better than the interrater agreement. The worst intrarater agreement was seen for the final DCIS group using the VN system.

Table 4 summarizes the proportions of slides with four or more nonexpert pathologists in agreement with the expert for final DCIS grade. Although the expert evaluating HL rated three slides as "special-type," the nonexpert pathologists were not given an opportunity to do so (according to the original classification publication criteria). Therefore, the data for these three slides were omitted for this classification. When the analysis is reduced to a common set of 37 slides that were reviewed by all, the overall majority agreement changes slightly. Agreement for HL remains significantly better than VN (p value moving from 0.02 to 0.04) but becomes nonsignificant compared with LA (p value moving from 0.04 to 0.06). There is no significant difference between LA and VN (p = 0.82).

Overall, the majority of NH pathologists agreed with the HL expert in 84% of cases compared with 63% and 60% agreement with the LA and VN experts, respectively. Using the VN system, the highest percentage of NH pathologists (93%) agreed with the expert in diagnosing a group 3 DCIS (high-grade nuclear features with or without necrosis), but the poor to fair kappa statistics for agreement with the expert for an overall final DCIS and overall final nuclear grade (kappa = 0.29 and 0.35, respectively) suggest that although the nonexpert pathologists distinguished well between the presence and absence of necrosis, agreement for nuclear grade (and hence overall final DCIS grade) was poor. Using HL and LA, the highest percentage of NH pathologists (92% and 73%, respectively) agreed with the experts in diagnosing an intermediate-grade DCIS. Although the moderate to good kappa statistics for agreement with the experts for an overall final DCIS and overall final nuclear grade (kappa = 0.53 and 0.49, respectively) reflect a comparable ability for the nonexpert pathologists to distinguish the cytonuclear and architectural differentiation of DCIS according to HL, this does not appear to be true for those using LA (kappa = 0.29 and 0.35, respectively).

When all assessments were considered for the final grade using all three classifications, only a small minority of the nonexpert interpretations (1.3%, 2.6%, and 4.6% for HL, LA, and VN, respectively) differed by two grades from the expert. A one-grade difference was observed in 31.8% (HL), 39.1% (LA), and 28.1% (VN) of cases. A general kappa provides an equal penalty when a nonexpert differs from the expert by either one or two final grade categories. We observed that the kappa values improved when weighted (less penalty was provided when a nonexpert differed from the expert by one final grade category compared with two) but there was no real qualitative difference.

## DISCUSSION

The most important predictors of clinical behavior for DCIS are the tumor histology, tumor size, and margin status.[33] Other criteria, of less proven prognostic significance, include the presence or absence of biologic markers such as hormone receptors[6] or metallothionein expression,[13] correlation with the mammographic findings,[17,21] cell kinetics,[19,25] and oncogene markers.[28,39] The clinical use of a histologic classification for DCIS depends on the reproducibility of its criteria.[3,5,22] Recent disagreement among pathologists regarding the diagnostic features of DCIS and predictors of local recurrence or invasive carcinoma have fueled the controversy regarding optimal therapy for DCIS.[16,31,33]

The traditional classification system for DCIS was based on architectural patterns and the presence or absence of necrosis.[1] Because more than one architectural pattern is often present in a single DCIS lesion, this criterion does not appear to be a reliable predictor of biologic behavior, and interrater reproducibility is poor.[12,37] In 1989, Lagios et al.,[22] using only nuclear grade and necrosis to classify DCIS, described a relationship between the tumor histology and risk of local recurrence in women choosing breast-conserving therapy. The recurrence rate was greater in cases of high-grade DCIS with comedo-type necrosis. Since 1994,

**TABLE 4.** *Proportions of slides with four or more pathologists in agreement with the expert for final DCIS grade*

| | Poorly differentiated/ high grade/group 3 % (frequency) | Intermediately differentiated/ intermediate grade/group 2 % (frequency) | Well differentiated/ low grade/group 1 % (frequency) | Special type | Overall % (frequency) |
|---|---|---|---|---|---|
| Holland (HL) | 76 (13/17) | 92 (11/12) | 88 (7/8) | * | 84 (31/37) |
| Modified Lagios (LA) | 63 (10/16) | 73 (8/11) | 50 (6/12) | 100 (1) | 63 (25/40) |
| Van Nuys (VN) | 93 (13/14) | 36 (4/11) | 47 (7/15) | — | 60 (24/40) |

* Although the Holland experts rated three slides as "special type," the non-expert pathologists were not given an opportunity to do so. Therefore, the data for these three slides were omitted for Holland.

Chi-squared p values for comparison of overall proportions: HL versus VN: p = 0.02, HL versus LA: p = 0.04 and VN versus LA: p = 0.82.

multiple classifications of DCIS have been proposed.[3,10,18,28,32,33,35] Most deemphasize the importance of architectural pattern but retain a three-tier system of final tumor grade (low, intermediate, high) and include criteria such as nuclear grade, necrosis, and cellular polarization.

The constant publication of new or modified classifications for DCIS presents a dilemma to many practicing pathologists who must decipher variable changes in criteria definitions and convey these changes in a meaningful and consistent way to their clinical colleagues. It also presents problems for researchers who are seeking to study treatment effectiveness based on prognostic factors. The authors of these classifications, usually with a special interest in breast pathology, may have documented diagnostic reproducibility data among themselves, but this does not necessarily translate into comparable agreement among pathologists with and without expertise in breast pathology.

A consensus classification for DCIS was published in 1997.[10] Although it is encouraging that so many eminent pathologists, surgeons, mammographers, radiation oncologists, and biostatisticians were willing to address collectively the defining features of DCIS and subsequent risk of local recurrence or invasive cancer, current reproducibility data for three of the four histologic criteria recommended for inclusion (nuclear grade, necrosis, polarization, and architectural pattern) either indicate poor agreement or have not been tested formally among nonexpert pathologists. The written criteria for nuclear grade put forward by the consensus classification appear to be a combination of those defined in the VN system[33] and the LA system.[32] Good interrater agreement for nuclear grade has been observed for both of these classifications.[4,32,34] However, other studies have indicated poor agreement in the reporting of necrosis and architectural pattern,[4,12,36,37] and cell polarization has only recently been evaluated separately for reproducibility.[15]

Our study assessed the diagnostic reproducibility and accuracy of three DCIS classifications by comparing the results of pathologists without expertise in breast pathology interpretation with those of three breast pathology experts, each of whom is a proponent of one of the three classifications evaluated. Prior reproducibility studies have usually involved pathologists with a special interest in breast pathology, and no other studies have made comparisons of the diagnoses with a reference standard. HL, a classification used by the European Pathologists Working Group, provided the best overall diagnostic accuracy and agreement among NH pathologists. The level of agreement among NH pathologists for HL (kappa = 0.46) was higher than that observed among 23 European pathologists who have a special interest in breast pathology for the same classification system (kappa = 0.37).[15] The European study recorded best overall agreement for the VN system (kappa = 0.42). This is of particular interest because HL, a familiar classification to European pathologists, is not used routinely in the United States whereas the VN system is. This finding may represent the increased concentration and attention required to review the set of slides according to an unfamiliar classification system. The best interrater diagnostic reproducibility for overall nuclear grade (cytonuclear differentiation) was seen in our study using HL (kappa = 0.45), which correlated with a moderate diagnostic agreement with the experts (kappa = 0.49).

The classification of Holland et al.[18] emphasizes the cytonuclear differentiation primarily, and the architectural differentiation (cellular polarization) secondarily. This classification has been found to correlate with oncogene and cell proliferation markers,[5,39] and mammographic microcalcifications,[17] and there appears to be a direct relationship between the grade of DCIS according to this classification system and the grade of invasive carcinoma.[23] A reproducibility study from New Zealand recorded improved agreement among the 11 participants compared with the traditional architectural classification, and most of the disagreements were in the distinction between the well- and intermediately differentiated groups.[4] These findings are also confirmed by Douglas–Jones et al.[12] The comprehensive evaluation of five DCIS classifications by 23 European pathologists (the European Commission Working Group on Breast Screening Pathology) found that the inclusion of cell polarization, as well as nuclear grade, in reaching a final DCIS grade using HL neither improved nor worsened the level of consistency that could be achieved using nuclear grade only.[15] In the current study, the reproducibility of cell polarization (kappa = 0.36, fair agreement) was less than that of nuclear grade (cytonuclear differentiation; kappa = 0.45, moderate agreement), but the comparable ability of the nonexpert pathologists to identify accurately poorly, intermediately, and well-differentiated final grades according to HL may reflect the influence of evaluating associated cell polarization. HL and LA assessed most accurately the intermediate final DCIS grades, and VN the high (group 3) final grades. These findings appear to reflect the ability of the NH pathologists to reproduce better the different grades of cytonuclear differentiation, rather than architectural patterns, according to HL, and the amounts of necrosis, rather than the cytologic characteristics, according to VN and LA.

The Van Nuys Prognostic Index,[33] developed to aid the process of optimal treatment selection, not only considers the histologic features of DCIS but also recognizes the importance of the lesion size and margin status. At the time of its publication, there were no studies validating the reproducibility of the histologic criteria. The technical problems in evaluating final margin status and tumor size in archival cases, such as multiplicity of bi-

opsies, multifocal lesions, standardized sampling of the entire lesion, and mammographic correlation, were also raised.[30] However, since then, studies have shown higher interrater agreement for this classification compared with others, as well as clinical correlations for grade of infiltrating carcinoma and disease-free survival.[4,12,15,34]

The accuracy and reproducibility for reporting necrosis in a two-tier system according to VN (present or absent) were high (kappa = 0.59 and 0.67, respectively) when compared with the three-tier system of LA (extensive, focal, or absent; kappa = 0.45 and 0.63 for accuracy and reproducibility, respectively). Because the published definitions for necrosis using these classifications and others are so variable, it is not surprising that there were high levels of agreement with the experts for reporting a presence or absence of necrosis. Agreement drops when pathologists are asked to categorize descriptive quantities of necrosis.

In the United Kingdom National External Quality Assessment scheme,[36] consistency of DCIS reporting was found to be good for the presence of a comedo growth pattern but poor for distinguishing architectural subtypes. A study comparing the VN and HL schemes found that, overall, the former was more reproducible, but that the evaluation of necrosis was an inconsistent criterion,[4] findings echoed by the European Commission Working Group on Breast Screening Pathology.[15] Neither of these studies assessed intrarater variability. A comparison of six classifications reported greatest disagreements for the assessment of architectural patterns and best agreement for the assessment of necrosis (present, absent, extensive).[12]

The most recently published histologic classification by Scott et al.,[32] using modifications of the criteria described initially by Lagios,[21] continues to support the clinical relevance of a three-tiered (final grades of high, intermediate, or low) rather than a two-tiered (comedo or noncomedo) system, with 94% agreement among six independent raters. Although the HL system does not attempt to include the comparatively rare special types of DCIS into its three-tier classification scheme, Scott et al.[32] have included a fourth option called "special-type" (such as pure apocrine or pure micropapillary DCIS types) in addition to a three-tier system of low, intermediate, and high grades. The inaccurate classification of these special-type lesions has often been documented.[3,26] Indeed, the only case that the expert using LA classified as special-type (apocrine) was not the same as the three cases that the expert using HL suggested were most likely an apocrine special-type DCIS. Other strengths of the LA classification are its known relevance to clinical outcome, its presentation as a reproducible system among its proponents, and the promise of an evaluation of interrater variation in a large multi-institutional study.[32]

For all three classifications, the intrarater reproducibility was better than interrater reproducibility, suggesting that each nonexpert pathologist had established his or her own fixed definitions of the criteria, even if these did not correlate well with the published criteria. This suggests that although formal study sets/tutorials before the study may increase agreement among nonexpert pathologists, the introduction of such study sets does not reflect either everyday practice of most pathologists or their ability to interpret new histologic criteria as they are published. However, it would be of interest, in a future study, to ask the same group of nonexpert pathologists to review a different set of slides after taking part in formal study groups detailing each classification, to assess any interpretation improvement.

Although studies comparing current DCIS classifications have indicated better reproducibility results for certain criteria, only one study has evaluated prediction of local tumor recurrence comparing multiple classifications.[2] This study showed a significant ($p = 0.009$) correlation between nuclear grade, as defined by the HL classification, and tumor recurrence when cell polarization was disregarded (a classification used currently by the United Kingdom National Health Service and European Commission-Based Breast Screening Programmes). A significant ($p = 0.001$) correlation between histology and recurrence was also observed using the VN classification with nuclear grade and necrosis. However, this study[2] did not control for adequacy of local excision because many of the archived cases had not been inked according to today's standard techniques.

The histologic classification of a DCIS lesion is only one factor to be considered when evaluating treatment options. The current study, and others, have addressed the diagnostic reproducibility of histologic criteria. However, an accurate, standardized determination of tumor size and margin status, requiring systematic and sequential processing of the tissue with mammographic correlation, is also important but rarely assessed formally.[20,33] In NH, a community-based quality improvement program in breast pathology addressed issues of tissue processing and standardization of surgical report content[7] to aid consistency of reporting. A national approach to such quality improvement issues may facilitate substantially the accuracy of pathology data assessed in multicenter trials.[27]

Three potential limitations of this study merit consideration. First, willingness to take part in such a slide review must be considered a potential bias in participant selection. The seven NH pathologists all took part in a prior statewide quality assurance study[38] and volunteered to take part in this study when its details were made known. These volunteers might have a greater interest in diagnostic reproducibility than the pathology community as a whole, and the participants almost certainly took

more time in the interpretation of these slides than for routine cases. Second, the participants were only asked to evaluate one slide from each case, acquired from the same institution. This may increase the potential for tissue sampling, processing, and staining variability interfering with slide interpretation. Indeed, the expert using HL cited tissue staining as the reason for favoring the diagnosis of apocrine, special-type DCIS in three cases. Third, the uniform reporting form may have influenced final interpretations because its format discouraged wordy comments and it may have differed from the reporting templates usually used by those pathologists.

In summary, the relevance and reproducibility of one published set of diagnostic criteria for classifying DCIS compared with another published set remains a serious issue. To ensure optimal treatment for a patient with DCIS, the histologic criteria should predict accurately tumor recurrence rates and should be reproducible diagnostically among pathologists without expertise in breast pathology interpretation. In the current study, HL provided the best overall diagnostic agreement with the experts and among NH pathologists. However, of the histologic criteria defined in all three of the tested classifications, nuclear grade (cytodifferentiation) according to HL and the presence or absence of necrosis were the best reproduced, but only fairly to moderately so. The importance of nuclear grade and necrosis in all proposed DCIS classifications is well known, but more research is needed to establish one internationally accepted set of simple and clear definitions for these criteria, with few subcategories, all tested for reproducibility among pathologists with and without expertise in breast pathology interpretation. This may be facilitated by slide study sets, photomicrographs, or digitized images on the Internet in addition to the information provided in the standard publication format. □

## Acknowledgments

The authors gratefully acknowledge the enthusiastic support of the NH Breast Pathology Quality Assurance Project Working Group (Alexander J. Bonica, MD; Constance A. Buttlar, MD; Dennis J. Carlson, MD; Jorge L. Gonzalez, MD; Michael D. Lagios, MD; Rosemary R. Millis, MB, BS, FRCPath; David L. Page, MD; Monira K. Rifaat, MD; Sonya M. Schwaegerle, MD; and Gary L. York, MD). They also extend their sincerest appreciation to Margaret Murphy at the New Hampshire Division of Public Health Services for her assistance in obtaining funding.

Supported by the New Hampshire Division of Public Health Services through a cooperative agreement with the Centers for Disease Control (grant no. U57-CCU108362-02) and by the United States Department of Defense (DAMD17-94-J-4109).

## REFERENCES

1. Azzopardi JG. Epitheliosis and in situ carcinoma. In: Bennington JL, ed. *Problems in Breast Pathology*. Philadelphia, PA: WB Saunders, 1979:128–46.
2. Badve S, A'hern RP, Ward AM, et al. Prediction of local recurrence of ductal carcinoma in situ of the breast using five histological classifications: a comparative study with long follow-up. *Hum Pathol* 1998;29:915–23.
3. Bellamy CO, McDonald C, Salter DM, Chetty U, Anderson TJ. Non-invasive ductal carcinoma of the breast. The relevance of histologic categorization. *Hum Pathol* 1993;24:16–23.
4. Bethwaite P, Smith N, Delahunt B, Kenwright D. Reproducibility of new classification schemes for the pathology of ductal carcinoma in situ of the breast. *J Clin Pathol* 1998;51:450–4.
5. Bobrow LG, Happerfield LC, Gregory WM, Springall RD, Millis RR. The classification of ductal carcinoma in situ and its association with biological markers. *Semin Diagn Pathol* 1994;11:199–207.
6. Bur ME, Zimarowski MJ, Schnitt SJ, Baker S, Lew R. Estrogen receptor immunohistochemistry in carcinoma in situ of the breast. *Cancer* 1992;69:1174–81.
7. Carney PA, Eliassen MS, Wells WA, Swartz WG. Can we improve breast pathology reporting practices? A community-based breast pathology quality improvement program in New Hampshire. *J Community Health* 1998;23:85–98.
8. Carney PA, Poplack SP, Wells WA, Littenberg B. The New Hampshire Mammography Network: the development and design of a population-based registry. *AJR Am J Roentgenol* 1996;167:367–72.
9. Choi WS, Parker BA, Pierce JP, Greenberg ER. Regional differences in the incidence and treatment of carcinoma in situ of the breast. *Cancer Epidemiol Biomarkers Prev* 1996;4:317–20.
10. Consensus Conference Committee. Consensus conference on the classification of ductal carcinoma in situ. *Cancer* 1997;80:1798–802.
11. DerSimmonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177–88.
12. Douglas-Jones AG, Gupta SK, Attanoos RL, Morgan JM, Mansel RE. A critical appraisal of six modern classifications of ductal carcinoma in situ of the breast (DCIS): correlation with grade and associated invasive carcinoma. *Histopathology* 1996;29:397–409.
13. Douglas-Jones AG, Schmid KW, Bier B, et al. Metallothionein expression in duct carcinoma in situ of the breast. *Hum Pathol* 1995;26:217–22.
14. Ernster VL, Barclay J, Kerlikowske K, Grady D, Henderson C. Incidence of and treatment for ductal carcinoma in situ of the breast. *JAMA* 1996;12:913–8.
15. European Commission Working Group on Breast Screening Pathology. Consistency achieved by 23 European pathologists in categorizing ductal carcinoma in situ of the breast using five classifications. *Hum Pathol* 1998;29:1056–62.
16. Fisher ER, Costantino J, Fisher B, Palekar AS, Redmond C, Eleftherios M. Pathologic findings from the National Surgical Adjuvant Breast Project (NSABP) protocol B-17: intraductal carcinoma (DCIS). *Cancer* 1995;75:1310–9.
17. Holland R, Hendriks JHCL. Microcalcifications associated with ductal carcinoma in situ: mammographic–pathologic correlation. *Semin Diagn Pathol* 1994;11:181–92.
18. Holland R, Peterse JL, Millis RR, et al. Ductal carcinoma in situ: a proposal for a new classification. *Semin Diagn Pathol* 1994;11:167–80.
19. Killeen JL, Namiki H. DNA analysis of ductal carcinoma in situ of the breast. A comparison with histological features. *Cancer* 1991;68:2602–7.
20. Lagios MD. Duct carcinoma in situ: biological implications for clinical practice. *Semin Oncol* 1996;23:6–11.
21. Lagios MD. Duct carcinoma in situ: pathology and treatment. *Surg Clin North Am* 1990;70:853–71.
22. Lagios MD, Margolin FR, Westdahl PR, Rose MR. Mammographically detected ductal carcinoma-in-situ: frequency of local recurrence following tylectomy and prognostic effect of nuclear grade on local recurrence. *Cancer* 1989;63:618–24.
23. Lampejo OT, Barnes DM, Smith P, Millis RR. Evaluation of infiltrating ductal carcinomas with a DCIS component: correlation of the histologic type of the in situ component with grade of the infiltrating component. *Semin Diagn Pathol* 1994;11:215–22.

24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
25. Meyer JS. Cell kinetics of histologic variants of in situ breast carcinoma. *Breast Cancer Res Treat* 1986;7:171–80.
26. O'Malley FP, Page DL, Nelson EH, Dupont WD. Ductal carcinoma in situ of breast with apocrine cytology: definition of a borderline category. *Hum Pathol* 1994;25:164–8.
27. Page DL, Lagios MD. Pathologic analysis of the National Surgical Adjuvant Breast Project (NSABP) B-17 trial [Editorial]. *Cancer* 1995;75:1219–22.
28. Poller DN, Silverstein MJ, Galea M, et al. Ductal carcinoma in situ of the breast: a proposal for a new simplified histological classification association between cellular proliferation and c-erbB-2 protein expression. *Mod Pathol* 1994;7:257–62.
29. Schnitt S, Connolly J, Tavassoli F, et al. Interobserver reproducibility in the diagnosis of ductal proliferative breast lesions using standardized criteria. *Am J Surg Pathol* 1992;16:1133–43.
30. Schnitt SJ, Harris JR, Smith BL. Developing a prognostic index for ductal carcinoma in situ of the breast: are we there yet? [Editorial] *Cancer* 1996;77:2189–92.
31. Schwartz GF, Finkel GC, Garcia JC, Patchefsky ASS. Subclinical ductal carcinoma in situ of the breast. Treatment by local excision and surveillance alone. *Cancer* 1992;70:2468–74.
32. Scott MA, Lagios MD, Axelsson K, Rogers LW, Anderson TJ, Page DL. Ductal carcinoma in situ of the breast: reproducibility of histological subtype analysis. *Hum Pathol* 1997;28:967–73.
33. Silverstein MJ, Lagios MD, Craig PH, et al. A prognostic index for ductal carcinoma-in-situ. *Cancer* 1996;77:2267–74.
34. Sneige N, Staerkel NP, Caraway NP, Fanning TV, Katz RL. A plea for uniform terminology and reporting of breast fine needle aspirates: the M.D. Anderson Cancer Center proposal. *Acta Cytol* 1994;38:971–2.
35. Tavassoli FA, Man Y. Morphofunctional features of intraductal hyperplasia, atypical intraductal hyperplasia, and various grades of intraductal carcinoma. *Breast J* 1995;1:155–62.
36. U.K. National Coordinating Group for Breast Screening Pathology, Sloane J, Ellman R, et al. Consistency of histological reporting of breast lesions detected by screening: findings of the U.K. National External Quality Assessment (EQA) scheme. *Eur J Cancer* 1994;30A:1414–9.
37. van Dongen JA, Holland R, Peterse JL. Ductal carcinoma in situ of the breast; second EORTC consensus meeting. *Eur J Cancer* 1992; 28:626–9.
38. Wells WA, Carney PA, Eliassen MS, Tosteson AN, Greenberg ER. Statewide study of diagnostic agreement in breast pathology. *J Natl Cancer Inst* 1998;90:142–5.
39. Zafrani B, Leroyer A, Fourquet A, et al. Mammographically-detected ductal in situ carcinoma of the breast analyzed with a new classification. A study of 127 cases: correlation with estrogen and progesterone receptors, p53 and c-erb-B2 proteins and proliferative activity. *Semin Diagn Pathol* 1994;11:208–14.

# APPENDIX D – CONFIDENTIALITY POLICY AND MANUAL

# CONFIDENTIALITY POLICIES AND PROCEDURES FOR DATA MANAGEMENT

New Hampshire Mammography Network
&
The Breast Cancer Surveillance Consortium

12/12/96

Submitted by

Patricia Carney, PhD
Dartmouth Medical School

Berta Geller, EdD
University of Vermont

We would like to acknowledge with great appreciation The University of Iowa and the University of California, San Francisco for the materials they shared, which contributed greatly to the development of this document.

## I. PURPOSE

This policy: 1) defines the types of confidential information collected, stored, utilized and transferred by National Cancer Institute Breast Cancer Surveillance Consortium members ; 2) outlines a minimal set of procedures for safeguarding this information; and 3) proposes an assignment of responsibilities within each contributing institution for these activities. The issue of protecting confidentiality in the use of patient and provider data is becoming increasingly more important as avenues for access, especially via computer, expand. The purpose of this policy is to provide a guide to Consortium members in data handling and use for maintaining confidentiality.

## II. BACKGROUND

The three major purposes of the National Cancer Institute Breast Cancer Surveillance Consortium are to 1) enhance our understanding of the operation and conduct of breast cancer screening in the United States, in part to respond to a congressional mandate in the Mammography Quality Standards Act (MQSA); 2) foster collaborative research among participants of the Surveillance Consortium to further our understanding of breast cancer screening; and 3) provide a foundation for the conduct of research, especially basic biological mechanistic research, aimed at improving our understanding of the natural history of breast cancer.

To achieve this purpose, each Consortium member site has established or is in the process of establishing a computerized registry of designated mammographic facilities within a specific geographic region. These registries have established or will establish linkage to the regional population-based cancer incidence registries or local cancer registry in order to assess various screening or diagnostic outcomes, such as the proportion of mammographic examinations that are abnormal, predictive value of mammography, and tests associated with follow-up of abnormal mammographic results in the community. Each Consortium member site collects, stores, utilizes, and transfers confidential data on mammography patients, physician's radiologic reports, and follow-up information, including pathology. These include clinical and epidemiologic data that are routinely collected on patients receiving mammography. These data are collected and may be used in collaboration with other investigators who may or may not be other Consortium members. The National Cancer Institute has funded a Statistical Coordinating Center, to which each site will be sending data for shared and pooled analyses. The term "Registry" will be used below to refer to any NCI Breast Cancer Surveillance Consortium member site.

## III. DEFINITION OF CONFIDENTIAL INFORMATION

Confidential information is information that contains identifying data, linking it to a specific research participant: patient, physician, or mammography practice. Such

identifying information includes, but is not limited to: patient, physician or facility name, address, telephone number, social security number, zip code, and/or occupation and employer. Confidential information also encompasses Registry proprietary information which includes, but is not limited to: copyrightable/patentable materials developed by Registry employees, consultants, and/or contractors.

Information generated by the Registry is classified into three categories based on the repercussions which may occur from unauthorized disclosure. These categories and their definitions are:

A. Public Information is information or data collected, compiled, utilized, or generated which is intended for public distribution and use or which may be obtained under freedom of information legislation. Generally, this includes aggregated data in published form, such as articles in medical journals about mammography patterns of care, accuracy, and other related topics. This does not include confidential information.

B. Internal Information is information or data collected, compiled, utilized, or generated by the organization which may be shared with employees and authorized consultants and contractors only. Authorization for external distribution or access shall be obtained from the Principal Investigator. Examples of internal information include mailing lists and technical proposals or software manuals.

C. Restricted Information is confidential information collected, complied, utilized, and/or stored by the organization which contains identifying links with specific individuals or medical practices such as name, address, or social security number. Confidential mammography registry data and reports fall within this category, as do any personal identifiers collected as part of Registry. Proprietary data or information produced by employees, consultants and/or contractors also falls within this category.

The Registry considers all data and information confidential that identify information specific to the patient, physician or facility specific information. Information that characterizes the case load of a specific institution or health care professional is considered proprietary and confidential.

## IV. THE RESPONSIBILITIES OF THE REGISTRY

The Registry's intent is to balance its research endeavors with its commitment to protect confidential information obtained and generated in the course of that research. It is the Registry's policy to adhere to laws and regulations that govern the collection, compilation, use, transfer, and storage of confidential data; to protect this information from unauthorized access or use at all time; to assure that this

information will only be transferred, utilized, and/or stored in sanctioned and approved ways; to assure that breaches of this policy are reported promptly and that appropriate corrective and/or disciplinary measures are taken; and to respond promptly to inquires from concerned participants regarding the Registry's research and other activities.

It is the responsibility of the Registry to protect the data from unauthorized access and release. The Registry maintains the same standards of confidentiality as customarily apply to the physician-patient relationship as well as the confidentiality of medical records. This obligation extends indefinitely, even after the patient is deceased or the physician ceases practicing within the area.

The costs of inappropriate release of confidential data are many. Inappropriate release of data could damage an individual whose diagnosis of cancer is made public; facilities and physicians could be severely compromised if accuracy or practice pattern data are disclosed. Legal protection of the data from discovery is necessary to assure that no harm comes to people contributing to the database.

- Government Regulations

Collection, access, use, and disclosure of confidential data pertaining to study subjects entered into the Registry and to proprietary information is governed by federal and state statutes and regulations. The Registry seeks to comply with these laws to the fullest extent possible to meet its obligations to funding sources and to meet its commitment to ethical principles upon which human subjects regulations are predicated.

1. State/Institutional Protection

Individual states may or may not have legislation in place that can provide protection from litigation to databases used for research purposes. If your state has this form of legislation, exploring whether the legislation has been tested in court will give you an indication of how advantageous it is likely to be in protecting research subjects. Quality assurance (QA) statutes have been used for years to protect participants contributing data to sensitive research projects. These institutional statutes are not as protective as they once were due to overuse. Because so many QA statutes have been overturned in court, they generally provide very little protection to databases or research subjects.

List and describe here all state laws, regulations and certificates pertaining to the registry:

---

---

---

---

2. The Federal Certificate of Confidentiality

The Breast Cancer Surveillance Consortium Members have applied for and received Federal Certificates of Confidentiality in accordance with the provisions of section 301(d) of the Public Health Service Act (42 U.S.C. 241 (d)). This certificate is issued to protect the privacy of research subjects by withholding their identities from all persons not connected with the research (See your site's certificate for the conditions that apply to the certificate).

3. Committees for the Protection of Human Subjects (CPHS)
Federal regulations guide institutional committees for the protection of human subjects. However, these regulations have various interpretations, depending on institution. Access to medical records via informed consent is becoming an increasingly controversial issue for institutional review boards. Working closely with your institution's CPHS in describing your project and ALL research subjects involved (providers as well as patients) will assist with compliance to these regulations and with the greatest level of protection by clearly identifying the research subjects.

## V. ACCESS TO THE DATA

- Registry Staff Members

Each staff member is required to read this Confidentiality Policy and Procedures Manual and signs a pledge to uphold this policy. The pledge remains in effect after cessation of employment. The Registry secretary (or personnel department) maintains a historical file of staff members who have signed pledges (See Appendix A for sample confidential agreement). The orientation and training of each new staff member shall include instructions concerning the confidentiality of data.

- Non-Registry Investigators and Other Interested Parties

Investigators or public health officials may request access to confidential or aggregate registry data. All requests shall be made in writing and approved by the Principal Investigator or an advisory body (such as a steering committee made up of community radiologists/pathologists and members of the Registry's research team). All procedures shall be followed and documented. All persons given access to data shall read the Confidentiality Policy and Procedures Manual and sign an agreement to adhere to the same confidentiality standards practiced by registry staff members. A formal data request form will be used for every request (See Appendix B for sample request form).

If an advisory committee is used, describe how the committee members are chosen or elected, their length of term and the procedures used to approve a request, including criteria; majority, unanimous, quorum etc.; time from request to approval; notification (See Appendix C for Sample Advisory Committee Operations Policy).
For data involving individual identifiers, requests shall be approved by an

approved Institutional Review Board (IRB) prior to submission of the request to the Registry.

Requests requiring the use of personal identifiers should indicate precautions for providing the necessary confidentiality in accordance with IRB standards, which includes reporting patient, practitioner and facility data in sufficient aggregate to minimize the risk of identifying individuals or individual practices. Any cells that have a small number of cases (which may identify an individual or a facility) shall be suppressed in those reports.

All requests shall clearly state the limits of data use. Data may only be used for the exact purpose for which they are requested. Data shall be kept confidential in the custody of the fewest individuals possible.

Data may only be released to the public for the purpose specified in the request. When data analyses are complete, data shall either be destroyed or, if needed for later reference, maintained in locked storage in the custody of an applicant for a specified period until they are no longer needed. Applicants shall specify the exact time period in their request during which they will require access to data.

All applicants shall agree to make a copy of any proposed publication or other form of public disclosure available to the registry 30 days prior to any public disclosure of data released from the registry. This will ensure adequate time to review, comment or decide to reanalyze and provide a response or alternate explanation, if necessary.

NOTE: FAILURE TO ABIDE BY TERMS OF THE AGREED USE OF DATA MAKES THE APPLICANT LIABLE FOR LITIGATION.

## VI. INAPPROPRIATE USES OF CONFIDENTIAL INFORMATION

Confidential data shall never be made available, to the extent allowed by law, for uses such as the following:

- Businesses that desire to market a product to patients.
- Health care institutions, their employees or providers that want to advertise or identify new patients for recruitment.
- Insurance companies that are attempting to determine the status of an individual.

## VII. DATA SECURITY

- General Data Management

The following components may be required to assure data security in all areas of Registry operation.

The Registry Director is ultimately responsible for data security.

Suitable locks are installed to control access to the Registry. Custodial staff are notified of the importance of maintaining a secure environment. A roster of persons authorized to enter the Registry is maintained by the Registry Administrative Personnel.

Registry staff are responsible for the confidentiality of all data encountered during data collection.

Confidential data shall not be transmitted from the Registry by any means (mail, telephone, electronic, or facsimile) without explicit authority from the Registry Director or a staff member to whom such authority has been delegated.

A registry-developed mail tracking system may be used to protect confidential data.

Precautions are taken for both physical and electronic security of confidential data sent on magnetic or electronic media.

Secure telephone data transmission includes using an unlisted telephone number, password access to the bulletin board systems, and restricted use of facsimile protect confidential data transmissions.

The physical security of confidential data stored on paper documents, computer printouts, microfiche and other media present in the Registry is ensured.

Confidential documents to be destroyed are kept in a secure environment until they are retrieved by a designated person or vendor for shredding and disposal.

- Report Handling

1. Physicians and Facilities Contributing Data to the Registry
For facilities that provide quality assurance data to contributing facilities/physicians, all physicians can receive reports on their own patients as per agreement with the Registry. These reports may contain identifying information indicating the radiologist or facility. Any report that contains patient level information shall be treated as confidentially as any medical record. For example, dummy codes can be generated each time a report is created to protect the identity of a receiving facility or radiologist. These codes shall never be able to link participants to actual study identifiers. Sites may also use a two step process for generating reports, where two individuals are responsible for report handling within a site, one will be kept blind to the dummy code, but will have access to the database for report production and one will be kept blind to the data source, but will apply the dummy code for processing and ultimate mailing. In generating reports requiring the use of personal identifiers, precautions for providing the necessary confidentiality in accordance

with IRB standards shall be undertaken. This includes reporting practitioner and/or patient data in sufficient aggregate to minimize the risk of identifying individuals or individual practice groups. Thus, any cells that have a small number of cases (which may identify an individual or a facility) shall be suppressed in those reports. Allowable uses of the report shall be clearly printed on the report or accompanying information. All requests for quality assurance data from other persons within the mammography facility shall have written approval from the physician or his/her designate physician in charge of quality assurance at said facility.

2. Contractor and Consultant Access
For those facilities who contract with computer programmers, biostatisticians etc., contractors and consultants who have access to restricted information shall read the Confidentiality Policy and Procedures Manual and sign a confidentiality agreement with assurances that they will safeguard such information from unauthorized access or further disclosure.

3. Statistical Coordinating Center (SCC)
A subset of the data collected at the Registry is transferred to the SCC of the National Cancer Institute's Breast Cancer Surveillance Consortium, located in Seattle, Washington. The data so transferred shall include no personal identifiers. As a member of the Breast Cancer Surveillance Consortium, the SCC has the same standards of confidentiality as all the member Registries.

- Procedure for Release of Data

Confidential data shall not be transmitted from the Registry by any means (mail, telephone, electronic, or facsimile) without explicit authority from the Principal Investigator or a staff member to whom such authority has been delegated. The specifics of the data (i.e. variables, date range) and to whom it will be transmitted shall be clearly communicated in writing to staff.

## VIII. COMPUTER SECURITY

Computers should be located in a locked facility which does not have public traffic. Computer security safeguards include the following:

Patient identifiers and demographic information are stored in files that have no other information. Other data are stored in separate computer files in the database. They are linked by a scrambled code that only authorized personnel understand.

There shall be password protection to enter Registry computers, applications and databases. All users accessing the database shall have a unique identification code and password. Passwords are changed on a regular basis or may be inactivated

if the users have not accessed them within a three month period. In this case, the individual would need to be reinstated to regain access to the databases. A user's identification and password shall be invalidated when the individual no longer requires access to the database.

All participating facilities and providers are given a confidential code number that is used in the database. A different confidential code number is assigned when reporting quality assurance data. This number is only known by appropriate staff, the facility and each individual provider.

Security standards strictly control access to the database files; staff have specific authorizations to read, write, erase or modify processed information.

Two copies of the daily backup shall be created. One back-up disk shall be stored in a locked file. The second backup disk shall be stored off site by an approved staff person. New staff shall be asked to store off-site backup disks after the probation period has ended. Registry backup disks should have no identification on them other than a number or code and a generic office address label. Caution will be taken to protect disks when off site by knowing where they are at all times and never leaving them in an unsecured location.

All word processing files that contain codes, passwords, data dictionaries or any descriptions of how to interpret the data should be stored in password protected files or removed from computers, copied onto disks/tapes and stored in locked cabinets.

An in-house printer should be used for the printing of confidential data, and the data never be left unattended in the printer.

Telephone data transmission are secured using an unlisted telephone number.

The use of personal and notebook computers for the ascertainment and management of confidential data is controlled by the same electronic and physical measures as described above to protect the security of the data.

Training and demonstration of computer systems are done with separate fictitious and/or anonymous data sets.

All disks/tapes containing Registry data shall be erased when not actively used for backup or transmitting of data.

- **Protection of Data and Network Connections at the SCC.**

1) Subject ID Encryption - All study identifiers at the site shall be recoded to a new SCC study identifier. To perform the recode, the SCC shall distribute a program based on a published algorithm (Meux, E Encrypting personal identifiers, Hlth Srvcs Res 1994, 29:247-256). The new SCC identifier cannot be reverse engineered to yield

the original identifier. The algorithm shall be used to recode subject identifiers, radiologist identifiers, and radiology site identifiers.. Only encrypted identifiers shall ever be sent to the SCC. All records sent to the SCC shall have the SCC identifier for internal record linkage.

2) Data Encryption - Data transmitted to the SCC shall be sent over the Internet, hence precautions shall be taken. Standard ASCII files (without variable identifiers) are encrypted using PKZIP and a password supplied to the site by the SCC. The encrypted data files are temporarily stored in the ftp area of mammstat.ghc.org. Within 24 hours the files are moved inside the GHC firewall to another computer. After the move the files are unencrypted.

3) Data Storage - The ftp area used by the SCC allows only the sites and NCI to logon. Once the files are moved to the computer inside the GHC firewall, only SCC staff shall have access to the data. The data are stored in Sybase with each file protected by a password. The data are resident only on a single computer and are not available on a network. To perform analyses, an analytic database is created that is then put on the network for use by the statistical analysts. Only analytic datasets shall be supplied to other users after approval by the publication committee.

APPENDIX A

## Sample Confidentiality Agreement

ACKNOWLEDGMENT:
I have read and agree to abide by the standards set forth in the CONFIDENTIAL POLICIES AND PROCEDURES FOR DATA MANAGEMENT, National Cancer Institute Breast Cancer Surveillance Consortium, _____, dated 10/10/96.

I am employed by: _____
(Name of Employer)

_____
(Address of Employer)

___ I am a consultant/contractor on:_____
(Name of Project)

___ I am a review committee member on: _____
(Name of Project)

___ I am an investigator requesting data for research.

___ I represent public health efforts and am requesting data.

___ I work at a mammography/pathology facility and request data for quality assurance purposes.

I understand that any confidentiality violations may make me liable for civil and/or criminal penalties.

DATE: _____

NAME:

_____
(Please Print)

_____
(Signature)

ADDRESS: _____
(Street) (City) (State/Zip)

## APPENDIX B

### Sample Confidential Data Request Form

Complete this form and return with any required documentation to:

Name: _____

Address: _____
_____
_____

Name of Applicant:_____

Institution: _____

Address: _____
_____
_____

Telephone: _____ FAX: _____ Date: _____

Title of Project: _____

Exact Data Requested:

_____
_____
_____

Exact purpose for which these data are being requested and limits of data use:
*Requests shall clearly state the limits of data use. Data may only be used for the exact purpose for which they are requested. Data may only be released for the purpose specified in the request.*

_____
_____
_____

Dates data use requested to begin __/__/__ ; to end __/__/__ .
*When data analyses are complete, data will either be destroyed or, if needed for later reference, maintained in locked storage in the custody of an applicant for a specified period until they are no longer needed.*
*Applicants will specify the exact time period in their request during which they will require access to data.*

PLEASE CONTINUE ON TO NEXT PAGE

Names and positions of persons responsible for maintaining data confidentiality *(Data shall be kept confidential in the custody of the fewest individuals possible; these individuals will sign a written assurance of confidentiality).*

Names Positions

---

---

---

---

---

---

For data involving individual identifiers, requests shall be approved by the Institutional Review Board (IRB) prior to submission of the request to the registry.

This request has received IRB approval dated: _ _/_ _/_ _
(or)
The request does not require IRB approval. _____ Initial here

*For requests requiring the use of personal identifiers, indicate precautions for providing the necessary confidentiality in accordance with IRB standards, which include reporting practitioner and/or patient data in sufficient aggregate to minimize the risk of identifying individuals or individual practice groups.*

Applicant agrees to make a copy of any proposed publication or other form of public disclosure available to the registry 30 days prior to any public disclosure of data released from the Registry.

_____ _____
Signature Date

Applicant shall cover the cost of retrieving data for this request to provide for use of the data without expense to the registry. Cost shall be determined by the Registry Director.

_____ _____
Signature Date

NOTE: FAILURE TO ABIDE BY TERMS OF THE AGREED USE OF DATA MAKES THE APPLICANT LIABLE FOR LITIGATION.

# APPENDIX C

## Sample Advisory Committee Policies

### New Hampshire Mammography Network - Guidelines for Advisory Committee

Selection of the N.H. Regional Breast Cancer Screening Network Advisory Committee members will be based on the following two criteria: 1) being a radiologist, mammography technologist, pathologist or researcher interested in and committed to the development of a mammography-pathology-tumor registry network that will enhance the quality of breast care in the state of New Hampshire and contribute to the study of breast cancer and breast cancer screening and, 2) having adequate geographic representation of mammographic centers state-wide.

Participation on this Committee will involve quarterly meetings. Attendance by conference call will be possible. The purpose of the Committee is to assist in the coordination and direction of efforts geared toward the implementation of the Department of Defense funded (DAMD17-94-J-4109) New Hampshire Regional Breast Cancer Screening Network. The primary responsibility of the Committee will be to determine policies and procedures that guide the conduct of this research. Membership terms will be reviewed annually.

The following are principles to follow and issues to consider regarding the collaborative efforts among members of the New Hampshire Regional Breast Cancer Screening Network Advisory Committee.

1) The Committee will keep in mind that the primary goals of this collaborative effort are to deepen our understanding of the practice of breast cancer screening and diagnosis in New Hampshire and elsewhere in the U.S., to further our understanding of breast cancer, and to produce high quality scientific work.

2) As its main functions, this Committee will help to develop the instruments needed for accurate data collection, and oversee the scientific activities and related analyses generated by the project. Members will be representatives of: New Hampshire radiologists and mammography technologists, the research team (including E. Robert Greenberg, MD, Patricia Carney, PhD, Steven Poplack, MD, Marguerite Stevens, PhD, Anna Tosteson, PhD), and a liaison from the state Health Department.

3) The Committee will meet quarterly for the first year of the project and semiannually thereafter for the remaining three project years.

4) Data Sharing:

a) As part of this project, data will be linked between the mammography and state tumor registries, both based in Hanover, New Hampshire and the New Hampshire State Department of Health and Vital Statistics, based in Concord, New Hampshire.

b) This project is part of the Breast Cancer Screening Surveillance Consortium, a consortium of eight federally-funded mammography programs, and it will contribute to both shared and pooled Consortium data. Pooled data is defined as analyses where site of origin or original population is disregarded. Shared data is defined as data from individual sites which may be analyzed and compared. Any decisions regarding data pooling and sharing will be made jointly by E. Robert Greenberg, M.D., Principal Investigator, and Patricia A. Carney, Director, and representative to the Breast Cancer Screening Surveillance Consortium Steering Committee with input from this Advisory Committee.

c) With the exception of contractual language (or grant language), data sharing will be done on a voluntary basis.

d) No identifying information will be part of any shared database.

5) Publications Policy:

a) A subcommittee of this Advisory Committee will sit as a publications advisory Committee.

b) A number of core analyses with the potential for turning into joint publications will be outlined by this Committee.

c) This Committee will draft a publications policy for the project and will establish a mechanism by which manuscripts can be shared among groups at the earliest appropriate time.

# APPENDIX E – PUBLICATION ON THE NCI BREAST CANCER SURVEILLANCE CONSORTIUM AND MEDICO-LEGAL ISSUES IN CONFIDENTIALITY AND DATA INTEGRITY

American Journal of Epidemiology

Vol. 152, No. 4
Printed in U.S.A.

# Current Medicolegal and Confidentiality Issues in Large, Multicenter Research Programs

Patricia A. Carney,[1] Berta M. Geller,[2] Howard Moffett,[3] Molly Ganger,[3] Matson Sewell,[1] William E. Barlow,[4] Nancy Stalnaker,[5] Stephen H. Taplin,[4] Cynthia Sisk,[4] Virginia L. Ernster,[6] Heather A. Wilkie,[6] Bonnie Yankaskas,[7] Steven P. Poplack,[1] Nicole Urban,[8] Michele M. West,[9] Robert D. Rosenberg,[10] Sharon Michael,[11] Thomas D. Mercurio,[12] and Rachel Ballard-Barbash[13]

The convenience of fast computers and the Internet have encouraged large collaborative research efforts by allowing transfers of data from multiple sites to a single data repository; however, standards for managing data security are needed to protect the confidentiality of participants. Through Dartmouth Medical School, in 1996–1998, the authors conducted a medicolegal analysis of federal laws, state statutes, and institutional policies in eight states and three different types of health care settings, which are part of a breast cancer surveillance consortium contributing data electronically to a centralized data repository. They learned that a variety of state and federal laws are available to protect confidentiality of professional and lay research participants. The strongest protection available is the Federal Certificate of Confidentiality, which supersedes state statutory protection, has been tested in court, and extends protection from forced disclosure (in litigation) to health care providers as well as patients. This paper describes the careful planning necessary to ensure adequate legal protection and data security, which must include a comprehensive understanding of state and federal protections applicable to medical research. Researchers must also develop rules or guidelines to ensure appropriate collection, use, and sharing of data. Finally, systems for the storage of both paper and electronic records must be as secure as possible. *Am J Epidemiol* 2000;152:371–8.

confidentiality; liability, legal; privacy

Information from medical records has contributed to data amassed in large databases for years. Cancer registries have been operating in many states for decades. More recently, other groups have designed and developed national or regional registries for childhood immunizations (1), cardiovascular surgery (2), and mammography screening (3–7). Health Employer Data Information Set performance measures derived from the databases of health maintenance organizations provide another example of computerized databases that contain potentially sensitive information. These databases are commonly aggregated for quality improvement or quality assurance purposes (2, 3, 8), as well as for research (9). Although confidentiality and integrity of data have always been a concern in research and clinical settings, technologic advances in data handling and the ability to share large data sets have made the process of protecting confidentiality more challenging. Potential harm to patients whose confidential medical information is disclosed ranges from invasion of privacy to potential exposure to exploitative marketing activities. This harm is widely recognized, and the legal mechanisms available to protect patients against such disclosure are fairly well understood. In contrast, potential harm to professionals (physicians, nurses, and other care providers), such as loss of economic security and vulnerability to litigation, is not as widely recognized or understood. What is clear, however, is that the overseers and users of confidential data must protect the interests of both patients and professional research participants.

The current literature on confidentiality lacks an outline of approaches to address relevant medicolegal issues for

Received for publication October 26, 1998, and accepted for publication October 11, 1999.

Abbreviations: IRB, institutional review board; NCI, National Cancer Institute; QA, quality assurance.

[1] Norris Cotton Cancer Center/Dartmouth-Hitchcock Medical Center/Department of Community and Family Medicine, Dartmouth Medical School, Hanover/Lebanon, NH.
[2] Office of Health Promotion Research, University of Vermont, College of Medicine, Burlington, VT.
[3] Orr and Reno, Professional Association, Concord, NH.
[4] Center for Health Studies, Group Health Cooperative of Puget Sound, Seattle, WA.
[5] Office of Sponsored Programs, University of Vermont, Burlington, VT.
[6] Department of Epidemiology and Biostatistics, University of California, San Francisco, CA.
[7] Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, NC.
[8] Fred Hutchinson Cancer Research Center, Seattle, WA.
[9] Department of Preventive Medicine, University of Iowa College of Medicine, Iowa City, IA.
[10] Department of Radiology, University of New Mexico, Albuquerque, NM.
[11] Cancer Prevention and Control Program, Colorado Department of Public Health and Environment, Denver, CO.
[12] General Counsel, University of Vermont, Burlington, VT.
[13] Applied Research Branch, Cancer Control Program, National Cancer Institute, Bethesda, MD.
Reprint requests to Dr. Patricia A. Carney, Norris Cotton Cancer Center, 1 Medical Center Drive, HB 7925, Lebanon, NH 03756-0001 (e-mail: Patricia.A.Carney@dartmouth.edu).

large computerized databases to which professional providers either knowingly or unknowingly have submitted information. We outline the approaches investigators should take to address data security and confidentiality for all research participants. This analysis is based on work conducted in eight states by investigators from seven academic research institutions, one health maintenance organization, and one state public health department. We analyzed federal and state laws as well as institutional policies intended to protect data from forced disclosure or use in litigation. We summarize the application of federal and state laws; describe essential steps for appropriate data collection, storage, utilization, and sharing; and offer confidentiality and security guidelines for data transfers from member sites to a central data depository. Our intention is to provide a clear framework for locally developed systems to protect the confidentiality of all research participants and ensure the integrity of data involved in confidential and sensitive medical research. It is critical that researchers carefully balance data use for the good of the public with a respect for the privacy and anonymity of all individuals involved.

## BACKGROUND

In 1994, the National Cancer Institute (NCI) funded a mechanism that would allow mammography registries to pool their data in one centralized database, in part to respond to a congressional mandate in the Mammography Quality Standards Act (10). The speed and efficiency this pooling allows enhances our understanding of the operation and conduct of breast cancer screening in the United States. The resulting collaborative, the National Cancer Institute Breast Cancer Surveillance Consortium, is described in detail elsewhere (11). Each Consortium member site had previously established a computerized registry that collects data from designated mammographic facilities within a specific geographic region. Each site sends its data electronically to a centralized Statistical Coordinating Center for pooled analyses. The data include confidential information on mammography patients, physicians' radiologic interpretive reports, and follow-up of abnormalities. Each mammography registry is linked to the regional population-based Surveillance, Epidemiology, and End Results registry of the NCI or similar statewide cancer registries. The linked data enable the determination of predictive value, sensitivity, and specificity of mammography as well as practice patterns. While these determinations are critically important to evaluating the performance of mammography, they necessarily involve the sharing of sensitive data, with potential risks to participants.

Soon after the Consortium was formed, a working group of representatives from each site obtained copies of federal and state laws that create a privilege against disclosure in litigation and of institutional regulations that address confidentiality of data generally. Our analysis of these materials revealed remarkable variability in how states address confidentiality issues. On the basis of our findings, we outline a recommended approach that investigators participating in large, multisite research programs may take in applying a minimum set of standards for the protection of all research subjects and health care providers and the data they contribute. In presenting this information, we will address definitions of confidentiality, the responsibilities of member sites, state and federal protections, data access, and paper and computer data security.

## DEFINITIONS OF CONFIDENTIAL INFORMATION

Confidential information is essentially personal information, that is, all information that links data to a specifically identified or identifiable research participant, professional or lay. Such identifying information may include physician (or patient) name, practice location name, address (including zip code), telephone number, occupation and employer, and, in some instances, rare diseases. Breach of confidentiality is the disclosure of health information without consent and without legal compulsion or legal authorization for its release (12). Table 1 outlines and defines categories of potentially sensitive information, ranked according to the severity of the potential repercussions of a breach of confidentiality.

## RESPONSIBILITIES OF MEMBER SITES

The objective of each member site is to balance its research endeavors with its commitment to protect confidential information obtained and generated in the course of that research. The policy of each site should be to adhere to laws and regulations that govern the collection, compilation, use, transfer, and storage of confidential data; to protect this information from unauthorized access or use at all times; to ensure that this information will only be transferred, utilized, and stored in sanctioned and approved ways; to ensure that breaches of this policy are reported promptly and that appropriate corrective or disciplinary measures are taken; and to respond promptly to inquiries from concerned participants regarding research and other activities. The obligation to protect data from unauthorized access and release extends indefinitely, even after the patient or physician is deceased or the physician ceases practicing within the area.

Adherence to applicable laws and regulations necessarily requires familiarity with the types of protections offered by federal and state governments and institutions. Table 2 outlines these types of protections, each of which is discussed more expansively in the next section. Table 3 outlines the types of protection available for each of the eight Breast Cancer Surveillance Consortium member sites.

## FEDERAL AND STATE LAWS AND REGULATIONS

The confidentiality of medical data is protected by laws and regulations at both state and federal levels. Collection, access, use, and disclosure of confidential data pertaining to study subjects entered at each member site are governed by federal and state statutes and regulations. In our medicolegal review, we focus on statutes and regulations protecting confidentiality. Although patients' privacy rights are recognized in medical ethics, common law, and constitutional law, statutes and regulations are the primary source of protection for research subjects. These sources also define parameters for use of medical records in research. Moreover, other significant confidential-

**TABLE 1. Categories and definitions of confidential information**

| Category of information | Definition |
|---|---|
| Public | Information or data collected, compiled, utilized, or generated that is intended for public distribution and use or that may be obtained under freedom of information legislation. Generally, this includes aggregated data in published form, such as articles in medical journals about patterns of care, accuracy, and other related topics. This does not include confidential information. |
| Internal | Information or data collected, compiled, utilized, or generated by the organization that may be shared with employees and authorized consultants and contractors only. Authorization for external distribution or access must be obtained from the Principal Investigator. Examples of internal information include site lists, technical reports, and research proposals in stages of preparation. |
| Restricted | Confidential information collected, compiled, utilized, and/or stored by the organization that contains identifying links with specific individuals or medical practices, such as name, address, and Social Security number. Confidential registry data and reports fall within this category, as do any personal identifiers collected as part of a registry (including diagnoses that, when linked with geographic location, could identify an individual or number of patients served by a facility that could identify provider participants). |

ity protections for patients, such as physician-patient privilege, are exclusively statutory creations. Each site must comply with these laws to the fullest extent possible to meet its obligations to funding sources and to meet its commitment to ethical principles upon which human subjects regulations are predicated. While federal laws are applicable to any state, state statutes, if they exist, can vary from state to state. The strongest possible legal protection exists where there are laws to protect confidentiality of data both from use in litigation (e.g., discovery or admissibility as evidence) and from forced disclosure of identifying information.

### Federal Certificates of Confidentiality

Federal Certificates of Confidentiality are issued in accordance with the provisions of section 301(d) of the Public Health Service Act, 42 U.S.C. section 241(d) to protect the privacy of research subjects by withholding their identities from all persons not connected with the research. Under Section 301(d), no federal, state, or local civil, criminal, administrative, legislative, or other proceedings can be used to compel disclosure of identifying characteristics of research subjects (13). This level of protection is the strongest and

**TABLE 2. Types of protection offered by federal or state governments and individual institutions**

| Type of protection |
|---|
| Federal |
| Public health service certificate of confidentiality |
| IRB* requirements for protection of subjects from the risk of loss of confidentiality |
| State |
| Laws protecting the confidentiality of records used in medical research |
| Laws protecting cancer or mammography registries |
| Quality assurance or peer-review statutes |
| Laws regulating physician-patient privilege |
| Laws on Patient's Bill of Rights |
| Laws governing confidentiality of patient's medical records |
| Institutional |
| Data security |
| Limiting data access with key or password protection |
| Outlining the specifics of all data handling using a standardized protocol |
| Shredding unneeded paper data |
| Formalizing all data requests and establishing a review process for release of research data |
| Developing a firewall for all computer systems |
| Maintaining off-site backups of computerized databases |
| Using a specially designed encryption program to convert data before sending it over the Internet |

* IRB, institutional review board.

TABLE 3. Types of protection available for each breast cancer surveillance consortium member site

| Consortium member states | Quality assurance statutes | Statutes establishing confidentiality protections for medical research | Mammography registry statutes | Cancer research/ registry statutes | Statutes addressing immunity for persons who furnish information to studies |
|---|---|---|---|---|---|
| California | X | X | | X | X |
| Colorado | X | X | | X | X |
| Iowa | X | X | | X | X |
| New Hampshire | X | X | | X | X |
| New Mexico | X | X | | | X |
| North Carolina | X | | | X | X |
| Washington | X | X | | X | |
| Vermont | X | | X | X | |

most comprehensive currently offered by applicable law, and legal precedent exists to support the strength of this coverage (14). The protection extends not just to patients and other research subjects, but also to professional participants (physicians, nurses, technologists, and other health care providers) who contribute data to each member site.

A decision to obtain a Federal Certificate of Confidentiality should be based on the potential risk of loss of confidentiality and a legal analysis of the level of protection offered by state statutes, which, as mentioned, is quite variable. The coverage afforded by the Certificates provides an important layer of uniform federal protection in addition to the variable protection offered at the state level and allows for protection of confidentiality of data crossing state lines, which is critical for sending data electronically (or otherwise) across state lines.

It is not necessary for research to be federally funded to be eligible for a Certificate of Confidentiality. However, Certificates are available only for research of a sensitive nature, such as research relating to sexual attitudes, preferences, or practices; use of alcohol, drugs, or other addictive products; illegal conduct; a situation that could, if released, be reasonably damaging to an individual's financial standing, employability, or reputation within the community; matters that would normally be included in a patient's record, disclosure of which could lead to social stigmatization or discrimination; or an individual's psychologic well-being or mental health (13). Additionally, applicants for a Certificate of Confidentiality must show that they will be engaging in a systematic study "directed toward new or fuller knowledge and understanding of the subject studied" (13, p. 729). Institutional Review Board (IRB) approval is required before an application for the Certificate is submitted. To cover professional participants, evidence of their status as research subjects must be provided, and the consequences of a breach of confidentiality must be specifically outlined. Information about the Certificate and application requirements can be obtained from either the NCI or the National Institute of Mental Health.

## State confidentiality laws

State laws protecting the confidentiality of records used in medical research can essentially be divided into five general categories: 1) laws specifically applicable to confidentiality of records used in medical research; 2) laws specifically applicable to cancer or other registries; 3) confidentiality requirements under quality-assurance or peer-review statutes; 4) laws creating a physician-patient privilege; and 5) laws generally applicable to the confidentiality of medical records. Protection afforded under all five types of legislation varies from state to state, although among most states, the first category consistently provides the most comprehensive protection for information collected for medical research. Considerations affecting coverage provided by each category of statute are briefly discussed below.

*Category 1. Medical research statutes.* Not all states have medical research statutes. In those that do, the adequacy of protection afforded depends upon several factors. We are unaware of any statute that specifically authorizes confidentiality protection for providers who are research subjects by virtue of reports or outcome data provided to the study, although in some states, statutory language may be expansively interpreted to provide that protection. Otherwise, the factors include whether confidentiality protection is needed for professional participants, whether the jurisdiction in which the research is conducted permits disclosure of information that identifies the participant as necessary to "further a study," and how personally identifying information is defined. Some statutes also prohibit redisclosure of information, while others are silent on this subject.

*Category 2. Registry statutes.* Some states have created programs for reporting incidences of disease to state registries. For research conducted pursuant to a state-authorized registry program, fairly strong confidentiality protection may be afforded by the applicable statute. These statutes often authorize disclosure of information collected by the registry to researchers, and researchers who work with such information may be entitled to confidentiality protection by the statute. Obviously, however, such laws are useful only for protecting the confidentiality of data collected in connection with a statutorily referenced registry.

*Category 3. Peer-review or quality-assurance (QA) statutes.* QA statutes and the scope of protection they afford differ widely from state to state. Although many researchers assume that QA statutes provide solid confidentiality protection, in fact, they often apply only to data col-

lected in very specific ways and for narrowly focused purposes. It may actually be possible to inadvertently waive the QA protection by using information collected for purposes that fall outside those authorized by the statute. Courts will likely find that QA statutes do not apply to protect the confidentiality of data if the following exist:

- Wrong class—The data are not within the class of information protected by the statute. Protected records typically include "records" generated by a quality-assurance, peer-review, or medical staff "committee." Researchers may not qualify as such a committee, and their research may not constitute a record.
- Formalities not observed—In some states, the committees must be a "regularly constituted review committee" of a hospital medical staff.
- Use for improvement of patient care—Many QA statutes protect only data that are systematically used to evaluate and improve the quality of patient care.
- Statutory exemption—An exception may apply to permit disclosure of information otherwise protected. For example, QA records usually can be disclosed in suits brought to challenge the denial of medical staff privileges.
- Absence of internal controls—Persons seeking to demonstrate that information is a protected QA record must usually be able to demonstrate that internal controls exist to protect the confidentiality of the subject data.
- Waiver—Confidentiality protection for QA records will be waived if otherwise protected information is voluntarily transferred outside the hands of the statutorily designated QA committee or office.

In summary and contrary to common perception, peer review or QA statutes may not confer substantial protection from discovery (15). The value of QA statutes in protecting the confidentiality of research databases is highly dependent upon how information is handled, by whom it is handled, and whether a legal precedent exists.

*Category 4. Physician-patient privilege laws.* Most states, if not all, have laws that establish an evidentiary privilege for communications between a physician and a patient about the patient's care. When the privilege applies, it prevents use of such communications in litigation. However, there are many exceptions to the privilege in most states. It is important to note that the privilege is generally said to "belong to the patient," meaning that only the patient (and not the provider) can claim it. As a result, the patient is free to authorize disclosure of the otherwise protected information to whomever he or she chooses. Because waiver of the privilege for one purpose may be held to constitute a waiver for other purposes, it is possible for patients to unwittingly authorize much broader disclosure than intended. The privilege may also be subject to statutory exceptions. Many states provide that it is inapplicable in proceedings before professional conduct committees. In sum, the privilege does not afford any protection to professional subjects of research, and the protection it gives patients may be quite limited.

*Category 5. Other laws generally applicable to the confidentiality of medical records.* Many states have adopted a Patient's Bill of Rights. These laws usually state that patients have the right to expect that communications and records pertaining to their care will be treated as confidential and not disclosed without their authorization. Privacy rights existing in the state and federal constitutions may also protect against disclosure of medical records in some instances. While these sources do not provide distinct protection for records collected by medical researchers, they may help bolster claims that medical information gathered by researchers is confidential. Because these laws change frequently, close surveillance is necessary by investigators who hope to access medical records for research purposes.

In addition to ensuring that the data are protected from legal discovery, researchers must be vigilant in protecting data from any use that might bring harm to the participants. This vigilance includes the establishment of both rules to prevent the misuse of data and systems to physically protect the data. These protections are discussed next.

## POLICIES AND PROCEDURES FOR HANDLING DATA

The orientation and training of staff members and investigators at all sites who require access to confidential data to conduct their work should include instructions concerning the collection, maintenance, use, and release of confidential data. Developing a policy and procedures manual brings a basic level of uniformity to data handling and access. Each new staff member should be required to read the confidentiality policy and procedures manual and sign a pledge to uphold this policy. The pledges must remain in effect after cessation of employment, so sites should maintain a historical file of staff members who have signed them.

At member sites, investigators or public health officials may request access to confidential or aggregate data. All such persons given access to data should read the confidentiality policy and procedures manual and sign an agreement to adhere to the same confidentiality standards practiced by the site's staff members.

Confidential data should not be transmitted from sites by any means (mail, telephone, electronic mail, or facsimile) without explicit authority from the Principal Investigator or a staff member to whom such authority has been delegated. The specific types of data, such as variables and date range, and those to whom they would be transmitted must be clearly communicated in writing to the staff. Because researchers often contract with computer programmers, biostatisticians, or contractors and consultants who have access to restricted information, these individuals should read the confidentiality policy and procedures manual and sign a confidentiality agreement with assurances that they will safeguard such information from unauthorized access or further disclosure. Confidential data should not be available to businesses or industries that desire to market a product or service to patients, health care providers or employees for advertising or recruitment of new patients, or insurance companies that are attempting to determine the status of individuals for any reason.

All external requests for data to be used in research should be approved by respective IRBs before submission of the request to the member site. All requests should be made in writing, preferably on a formal data request form, and should clearly state the limits of data use. Data may be used only for the exact purpose for which they are requested, must be kept confidential, and must remain in the custody of the fewest individuals possible. Applicants should specify the exact time period during which they will require access to data and should agree to provide a copy of any proposed publication or other form of public disclosure to member sites at least 30 days before release. This period will ensure adequate time to review, comment, or decide to reanalyze and provide a response or alternate explanation, if necessary.

All requests should be approved by the Principal Investigator or an advisory body, such as a steering committee made up of community physicians and members of the site's research team. If an advisory committee is used, a description of how the committee members are chosen or elected, their length of term, and the procedures used to approve a request should be outlined, including voting criteria (majority, unanimous, quorum), time limits for responding to requests for approval, and notification and documentation requirements.

Requests requiring the use of personal identifiers should explain the necessary precautions to be taken to provide confidentiality in accordance with procedures approved by the project's IRB, such as reporting patient, practitioner, and practice site data in sufficient aggregate to minimize the risk of identifying individuals or individual practices. When data analyses are complete, data should either be destroyed or, if needed for later reference, maintained in locked storage in the custody of an applicant for a specified period until they are no longer needed. If a central data repository is used for pooled analyses, this repository should abide by the same standards of confidentiality as all member sites. In addition, a review process for requests of pooled data should be developed.

## DATA SECURITY

### Paper systems

The following components can enhance data security in all areas of member site operation. Suitable locks should be installed to control access to the site, and all staff should be notified of the importance of maintaining a secure environment. A roster of persons authorized to enter the area should be maintained by the administrative personnel. Staff should be responsible for the confidentiality of all data encountered during data collection.

A site-developed mail-tracking system should be used to protect confidential data. The physical security of confidential data stored on paper documents, computer printouts, microfiche, and other media forms from member sites should be ensured. Confidential documents to be destroyed should be kept in a secure environment until they are shredded and disposed of properly.

If member sites produce QA reports for practitioners or other facilities at designated intervals, those receiving the reports should be informed about appropriate and inappropriate methods of handling them and should comply with applicable QA statutes. While legal protection from discovery is necessary to ensure that no harm comes to those contributing data to a database, the same individuals have an equal responsibility to protect the confidentiality of data they receive from member sites.

QA reports may contain identifying information about providers or patients. Any report that contains identifiable information must be treated as confidentially as any medical record. Encrypted codes may be generated when appropriate each time a report is created to protect the identity of a receiving practice location or radiologist. These codes should never link participant identifiers to actual study data. To provide extra protection when preparing report mailings, a two-step process may be used. Here, two individuals are responsible for report handling within a site, with one kept blind to the encrypted code and having access to the database for report production while the other, who applies the encrypted code for processing and ultimate mailing, is kept blind to the report content. Practitioner or patient data should be reported only in aggregate sufficient to minimize the risk of identifying individuals or individual practice groups. Thus, any cells that have a small number of cases (which may identify an individual or a practice location) should be suppressed in those reports. The purpose of the reports should be clearly printed on them or on accompanying information.

### Computer systems

Computers should be located in a locked facility with no access to public traffic. Computer security safeguards are outlined below.

1. Participant identifiers and demographic information should be stored in files that contain no other information. Other data should be stored in separate computer files in the database. They should be linked by a scrambled code that can be accessed only by authorized personnel.
2. Password protection should be required for the computers, applications, and databases of each member site. All users accessing the database should have a unique identification code and password. Passwords should be changed on a regular basis. A user's identification and password should be invalidated when the individual no longer requires access to the database. Precautions should be taken for both physical and electronic security of confidential data sent on magnetic or electronic media. Secure telephone data transmission should be accomplished by using an unlisted telephone number, password access to the bulletin board systems, and restricted use of facsimile technology for the transmission of confidential data.
3. Backup disks or tapes should have no identification on them other than numbers or codes and a generic office

address label. They should never be left in an unsecured location.

4. All word processing files that contain codes, passwords, data dictionaries, or any descriptions of how to interpret the data should be stored in password-protected files or removed from computers, copied onto disks or tapes on a weekly basis, and stored in locked cabinets. An in-house printer should be used for the printing of confidential data, which should never be left unattended in the printer.
5. The use of personal and notebook computers for the ascertainment and management of confidential data should be controlled by the same electronic and physical measures as described previously.
6. Training and demonstration of computer systems should be performed with separate fictitious or anonymous data sets.
7. All disks and tapes containing member site individual or pooled data should be erased when not actively used for backup or transmission of data.
8. When the site provides aggregate data to a centralized location, all study identifiers from the original site should be recoded to a new centralized study identifier. Performing the recode can be based on a published algorithm (16). It should not be possible to reverse engineer the new centralized identifier to yield the original identifier. The algorithm should be used to recode all identifiers. Only encrypted identifiers should be sent to centralized databank, all of which should have the centralized identifier for internal record linkage of longitudinal data.
9. Data transmitted to a centralized location can be sent over the Internet if precautions are taken. Standard ASCII files (without variable identifiers) should be encrypted using a special program and a password supplied to the site by the central program office. The encrypted data files should be temporarily stored in the file transfer protocol area of a centralized computer designated to receive data from the Internet. Within 24 hours, the files should be moved inside firewall protection to another computer. After this move, the data fields of the files can be unencrypted.
10. The file transfer process area used by the central program office should allow only member sites to log on. Once files are moved to a computer inside firewall protection, only centralized staff should have access to the data. The data should be stored in a master relational database, with each file protected by a password. The data should be available only on a private internal network accessed only by centralized statistical personnel with no Internet access. Only analytic data sets should be supplied to other users and only after approval by a steering committee or other governing body.

## DISCUSSION

The critical challenge in database research is to maintain the balance between the conduct of research for the good of the public's health and the protection of an individual's right to privacy. Large, multisite database studies, such as those being conducted by the NCI Breast Cancer Surveillance Consortium, can provide important data for shared or pooled analyses critical to addressing important public health issues. The major risk to participants is disclosure of potentially sensitive information and loss of confidentiality of identifying information. We worked collaboratively as a Consortium with legal consultation to identify, analyze, and outline how best the nine partnership sites could protect the confidentiality and integrity of data and databases. Our efforts identified several issues that deserve further discussion.

Although state QA laws can both prevent the release of individual-level information and protect data from use in litigation (17–19), care must be taken to comply with these laws and protection may be threatened by misuse of data (20, 21). Institutions and individual practitioners have relied on the QA or peer-review statutes in their respective states to confer protection from discovery for a variety of review and clinical improvement activities. In many instances, the protection, in fact, never existed, due to the manner in which information was gathered and processed and the results were distributed. To maintain protection, sites must gather and handle the information in a manner specified by the applicable state statute. It may not be possible to bring multifacility or multistate research projects into compliance with the QA laws; thus, it may be necessary to look for other sources of protection, such as a Certificate of Confidentiality.

Most states have laws that provide varying degrees of confidentiality protection to different kinds of medical records. However, the differences in the applicability of these laws can be significant. This issue is becoming increasingly controversial (21, 22), as the public has become more aware of occurrences of medical record misuse, including sales of medical records and release of medical information to federal program auditors and mortgage holders (20). National legislative activity has increased significantly in this area. On the national level, a comprehensive federal policy on confidentiality of medical records can be expected in the year 2000. The United States Congress has considered at least two recent legislative proposals that deal directly with attempts to ensure privacy of identifiable health information, such as the medical record (20). Issues concerning informed consent, disclosure, and physical security, as well as who would be the oversight body, are under consideration.

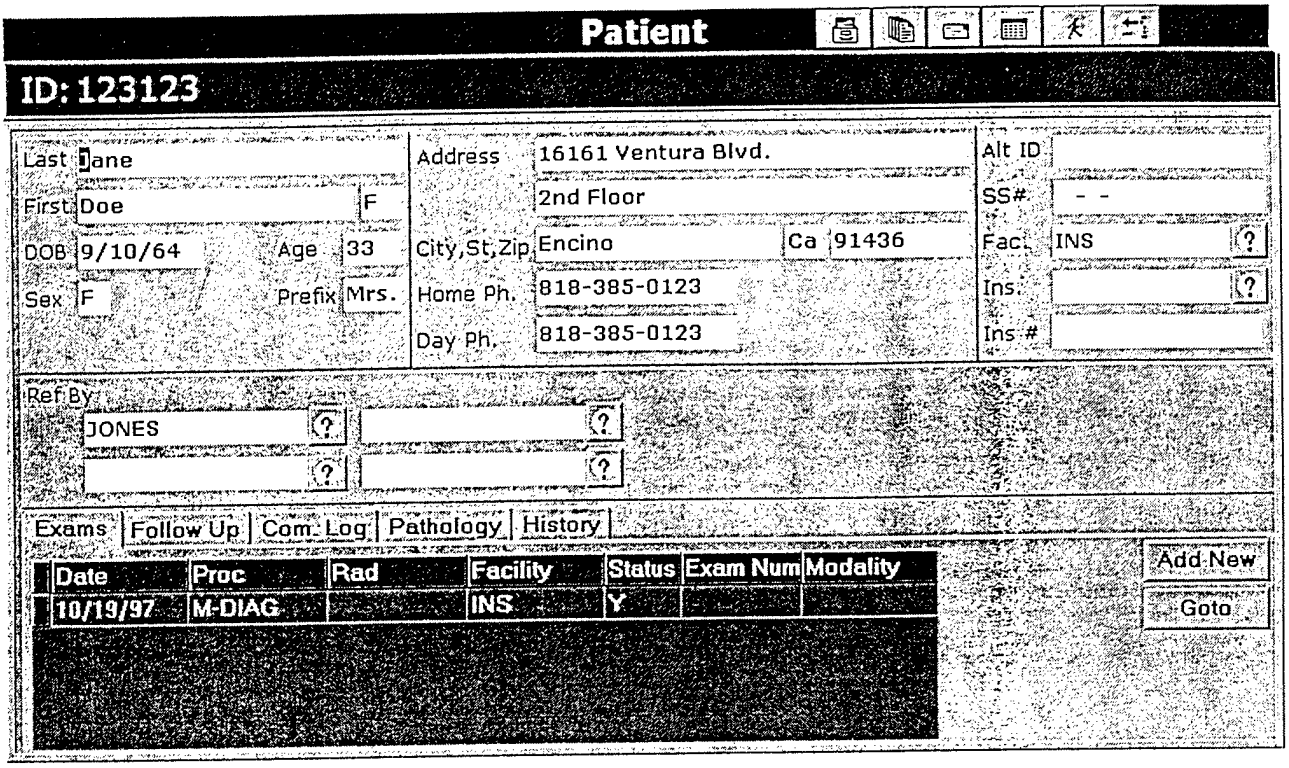
It is important for the public to understand and recognize the difference between utilizing medical information for the good of the public, such as is done in medical research, and medical record misuse that occurs outside the protection of the federal and state regulations discussed in this paper. For research studies to gain the participation needed by the public, the confidentiality of research data must be honored and protected. Otherwise, it will be impossible to conduct research such as that being done by our Consortium. It is equally important for researchers who intend to collect data for research purposes to rely on current laws and to monitor pending legislation that may affect their ability to conduct their research. The strongest legal mechanism of protection that currently exists is the Federal Certificate of Confidentiality. Its strength lies in the geographic coverage

it affords, the relative paucity of exceptions to its coverage, and the legal precedent that already exists regarding its use to protect the confidentiality of research subjects. Notwithstanding this valuable mechanism, researchers should be familiar with the specific confidentiality and privacy protections that may exist within their own jurisdictions and apply them when appropriate. To maximize protection, researchers should obtain a Certificate of Confidentiality; research legal precedents in their state and take advantage of the protection available; and institute measures to minimize the chance of unauthorized or inadvertent disclosure of confidential information in databases, data reports, and research information. Through these actions, researchers can fulfill their ethical and legal obligations by protecting confidential information to the maximum extent possible under existing law, while continuing their research.

## REFERENCES


1. Gostin LO, Lazzarini Z. Childhood immunization registries. A national review of public health information systems and the protection of privacy. JAMA 1995;274:1793–9.
2. Malenka DJ, O'Connor GT. A regional collaborative effort for CQI in cardiovascular disease. Northern New England Cardiovascular Study Group. Jt Comm J Qual Improv 1995; 11:627–33.
3. Smith R, Osuch JR, Linver MN. A national breast cancer database. Radiol Clin North Am 1995;33:1247–57.
4. Rosenberg RD, Lando JF, Hunt WC, et al. The New Mexico Mammography Project: screening mammography performance in Albuquerque, New Mexico, 1991 to 1993. Cancer 1996;78:1731–9.
5. Carney PA, Poplack SP, Wells WA, et al. The New Hampshire Mammography Network: development and design of a population-based registry. AJR Am J Roentgenol 1996;167: 367–72.
6. Geller BM, Worden JK, Ashley JA, et al. Multipurpose Statewide Breast Cancer Surveillance System: the Vermont experience. J Registry Management 1996;23:168–74.
7. Yankaskas BC, Jones MB, Aldrich TE. The Carolina Mammography Registry: a population-based mammography and cancer surveillance project. J Registry Management 1996; 23:175–8.
8. Donaldson MS. Gearing up for health data in the information age. Jt Comm J Qual Improv 1994;20:202–7.
9. Taplin SH, Mandelson MT, Anderman C, et al. Mammography diffusion and trends in late-stage breast cancer: evaluating outcomes in a population. Cancer Epidemiol Boio Prev 1997;6: 625–31.
10. Federal Register, Department of Health and Human Services, Food and Drug Administration, 21 CFR Parts 16 and 900. Volume 62, no. 208. Quality mammography standards—final rule. Docket 95-N-0192. October 28,1997.
11. Ballard-Barbash R, Taplin SH, Yankaskas BC, et al. Breast cancer surveillance consortium: A national mammography screening and outcomes database. AJR Am J Roentgenol 1997; 169:1001–8.
12. Turkington RC. Medical record confidentiality law, scientific research, and data collection in the information age. J Law Med Ethics 1997;25:113–29.
13. Early CL, Strong LC. Certificates of confidentiality: a valuable tool for protecting genetic data. Am J Hum Genet 1995;57: 727–31.
14. People v. Newman, 32 N.Y.2d 379, cert. denied, 414 U.S. 1163 (1973).
15. American Medical Association. A compendium of state peer review immunity laws. Chicago, IL: American Medical Association, 1988.
16. Meux E. Encrypting personal identifiers. Health Serv Res 1994;29:247–56.
17. Gostin LO, Lazzarini Z, Nesland VS, et al. The public health information infrastructure. A national review of the law on health information privacy. JAMA 1996;275:1921–7.
18. Bellach B, Schon D. Legislation to protect individual confidentiality: the case of cancer registration in Germany. Sci Total Environ 1996;184:33–6.
19. Irvine D. Confidentiality: data and permissible disclosure. J R Soc Med 1996;87 (Suppl. 22):42–3.
20. Marwick, C. Increasing use of computerized record keeping leads to legislative proposals for medical privacy. JAMA 1996;276:270–2.
21. Galbraith RA, Sobel BE. Patient information and confidentiality: a compelling conundrum. (Editorial). Coron Artery Dis 1997;8:59–60.
22. Gostin L. Health care information and the protection of personal privacy: ethical and legal considerations. Ann Intern Med 1997;127:683–90.

# APPENDIX F – INSIGHT DATA COLLECTION SCREENS


Patient
ID: 123123
Last Jane
First Doe
F
DOB 9/10/64
Age 33
Sex F
Prefix Mrs.
Address 16161 Ventura Blvd.
2nd Floor
City,St,Zip Encino
Ca 91436
Home Ph. 818-385-0123
Day Ph. 818-385-0123
Alt ID
SS# - -
Fac. INS
Ins.
Ins #
Ref By
JONES
Exams
Follow Up
Com Log
Pathology
History
Date
Proc
Rad
Facility
Status
Exam Num
Modality
10/19/97
M-DIAG
INS
Y
Add New
Goto

# Patient Intake (Tech) form

Patient Has read and Signed NHMN Survey Consent Form
○ Yes ○ No

Patient Has Had Previous Mammogram
○ Yes ○ No

Patient Has Breast Concerns
○ Yes ○ No

Who first became concemed?
○ Self ○ Partner ○ Physician/Nurse

How Long has there been Concern?

Had Breast Cancer
○ Yes ○ No

Age ○ L ○ R ○ B

Family History of Breast Cancer?
○ No ○ Yes ○ Unknown

Specify
○ Mother ○ Sister(s)
○ Other ○ Daughter(s

How many?

Have Periods stopped permanently?
○ No ○ Yes ○ Not Sure

Taking Hormones
○ Yes ○ No

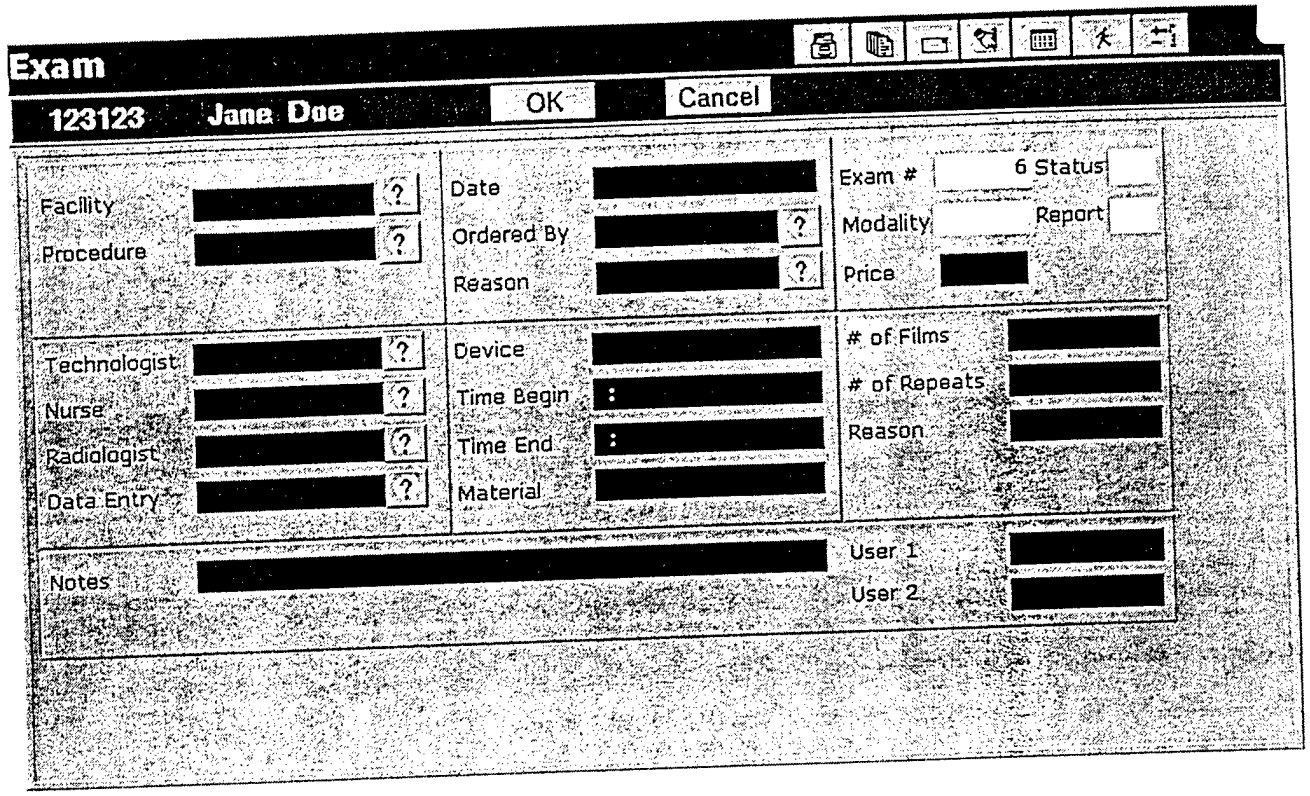
How Long?

Date | Location

| Type of Concern | L | R | B |
|---|---|---|---|
| Lump | ○ | ○ | ○ |
| Nipple Discharge | ○ | ○ | ○ |
| Skin Changes | ○ | ○ | ○ |
| Other (specify) | ○ | ○ | ○ |

Past breast procedures
○ Yes ○ No

| | L | R | B | Date |
|---|---|---|---|---|
| Breast Reduction | ○ | ○ | ○ | |
| Breast Implants | ○ | ○ | ○ | |
| Needle Biopsy | ○ | ○ | ○ | |
| Surgical Biopsy | ○ | ○ | ○ | |
| Lumpectomy | ○ | ○ | ○ | |
| Mastectomy | ○ | ○ | ○ | |
| Breast Reconstruction | ○ | ○ | ○ | |
| Radiation Therapy | ○ | ○ | ○ | |


Exam
123123 Jane Doe
OK
Cancel
Facility
Procedure
Date
Ordered By
Reason
Exam #
6 Status
Modality
Report
Price
Technologist
Nurse
Radiologist
Data Entry
Device
Time Begin
Time End
Material
# of Films
# of Repeats
Reason
Notes
User 1
User 2

# Exam Results

OK Cancel

Comparison Mammograms used for Interpretation
- Yes
- No

Breast Ultrasound used for Interpretation
- Yes
- No

Breast Composition
- Fat
- Scattered
- Heterogenously Dense
- Extremely Dense

**Assessment Status: (Choose 1 per Breast)**

Negative (ACR 1)
- B
- L
- R

- (ACR 0) Assessment Incomplete
- (ACR 2) Benign Finding-Negative
- (ACR 3) Probably Benign Finding
- (ACR 4) Suspicious Abnormality
- (ACR 5) Highly Suggestive of Malignancy

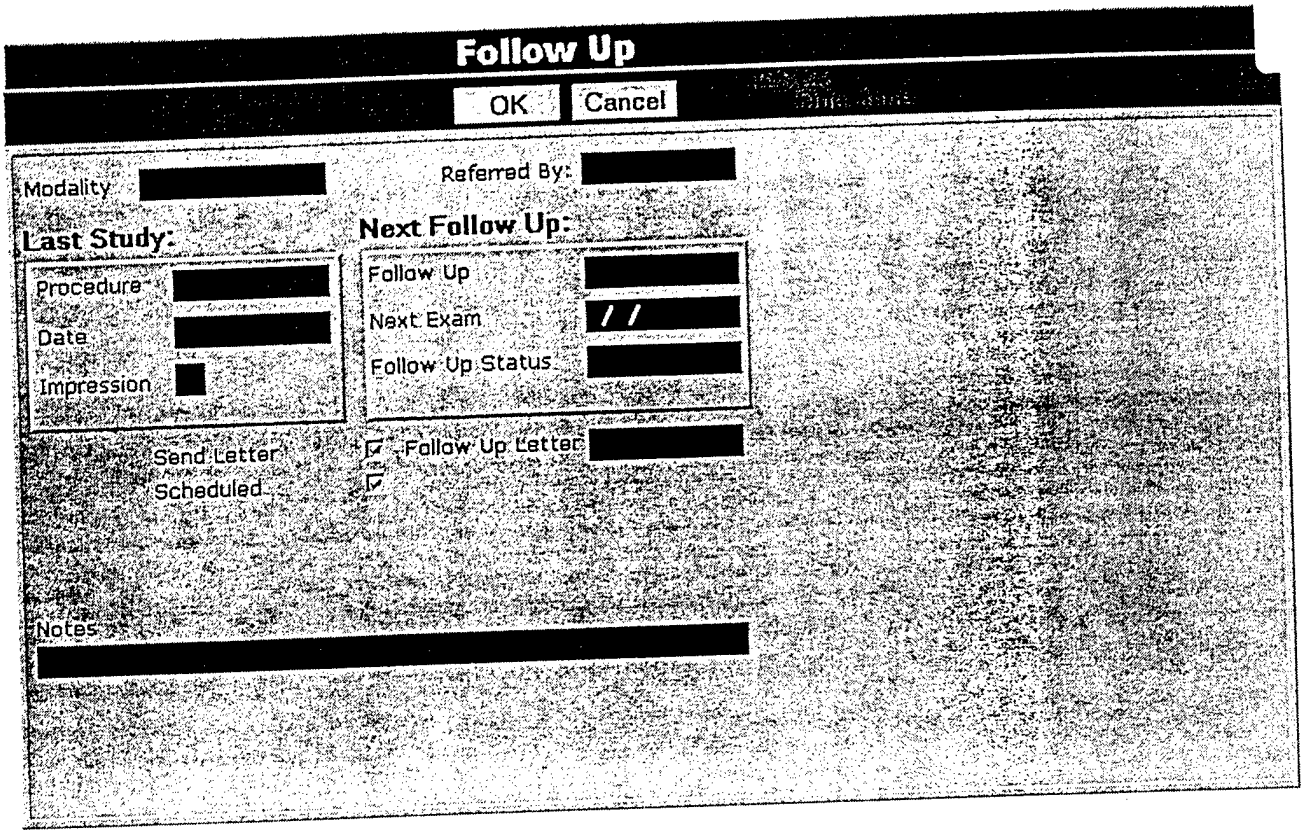
| | L | R | B |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

**Recommendation:**

Routine Screening Mammogram
- B
- L
- R

- Follow-Up Mammogram at Short Interval
- Additional Views to Supplement Current Exam
- Breast Ultrasound
- Clinical Breast Exam
- Surgical Consult
- Biopsy (including FNA)

| | L | R | B | Months |
|---|---|---|---|---|
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |


Follow Up
OK
Cancel
Modality
Referred By:
Last Study:
Procedure
Date
Impression
Next Follow Up:
Follow Up
Next Exam
/ /
Follow Up Status
Send Letter
Follow Up Letter
Scheduled
Notes

# APPENDIX G – STUDY PAPER DATA COLLECTION INSTRUMENTS

| MM | DD | YY |
|---|---|---|

# NH Mammography Network General Information

Patient's Name: _____
Last First Middle

Address: _____

_____

_____

Zip code: _ _ _ _ _

Today's Date: _ _ - _ _ - _ _
month day year

PLEASE CLEARLY PRINT ALL INFORMATION

**Please read the information below before you fill out the attached survey.**

## Information about the New Hampshire Mammography Network Project

Your mammography center is working with the Norris Cotton Cancer Center and Dartmouth Medical School to develop a registry (a computer database) of mammograms that will help us understand breast problems, including breast cancer. The registry is called the New Hampshire Mammography Network.

We are asking you to help us expand the usefulness of this registry by giving us additional information on the attached survey. The survey is for research purposes only. It is not part of your routine procedure for mammography. **Your participation is strictly voluntary. Whether you participate or not, your decision will have no effect on your medical care.**

The information you give us on the attached survey will be entered into our New Hampshire Mammography Network, along with your mammography results. However, if you are a resident of Vermont, your information will be transferred to a similar registry in Vermont. Neither our registry nor the Vermont registry will release any information that allows you to be identified. Although data collected may be shared with other investigators, your name and other identifying information will not be revealed.

If, after your mammogram, you have additional diagnostic studies or treatment related to breast problems, we may need to review your medical records to help us fully understand your mammography results. **In addition, pathology reports related to your breast diagnostic/procedure may also be requested. Rarely, we also may wish to contact a patient or her doctor directly to ask for more information. This may occur once or twice for every 200 mammograms we receive.**

**Please Note:** If there are any questions on the survey that you do not wish to answer, simply leave them blank. If you do not wish to participate in this research study, please hand all the forms back to the receptionist or mammography technologist. If you agree to participate, we will continuously include your mammography data in our study as you receive other mammograms. If at some point you wish to withdraw, please notify the receptionist or technologist.

If you have any questions regarding the NH Mammography Network Project, please call the Norris Cotton Cancer Center at 603-650-3414. Ask to speak with Martha Goodrich or Patricia Carney.

**Permission:** We ask your permission to use your data in our project, and, if needed, to review your record or to contact you or your doctor for additional information. Please sign here to indicate that you are willing to participate fully in these activities.

**Signature:** _____

*Thank you for your cooperation!*

*revised Exeter final 6/1999*

48440

# Patient Intake (Tech.) Form

B 0 0 9 0 2 1 4

DATA LINKS

Name: _____ (Last / First / Middle Initial)

Date of Exam: ____ - ____ - ____ (mm / dd / yy)

Social Security #: _____ - _____ - _____

Zip Code: _____

Date of Birth: ____ - ____ - ____ (mm / dd / yy)

Tech Initials: _____

Referring Physician's Name: _____

Referring Physician's Town: _____

Medical Record #: _____

**Did the Patient read & sign the NHMN Survey Consent Form?**

○ No ○ Yes

**Has the Patient had a previous mammogram?**

○ No ○ Yes

Date of Last Mammogram: ☐☐ / ☐☐ / ☐☐ (m m / d d / y y)

**Location/State:** _____

**Does the Patient have any breast concerns?**

○ No ○ Yes

If Yes, who **first** became concerned? (choose ONE)

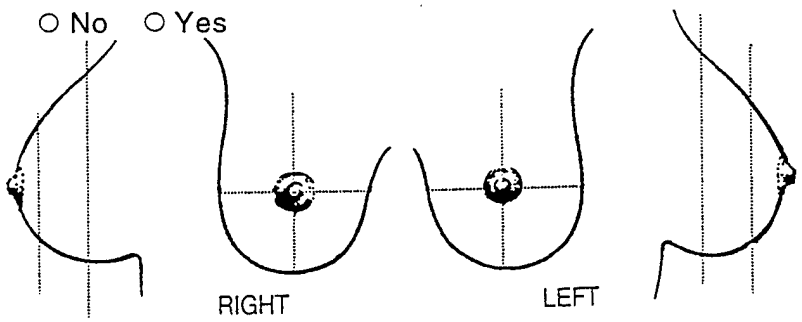
○ Self ○ Partner ○ Physician/Nurse

How long has there been concern? (e.g enter 01 for 1 month or less) ☐☐ Months

| Type of concern: | L | R | B | |
|---|---|---|---|---|
| Lump | ○ | ○ | ○ | |
| Nipple Discharge | ○ | ○ | ○ | |
| Skin Changes | ○ | ○ | ○ | |
| Other (please specify) | ○ | ○ | ○ | _____ |

**Has the Patient had any past breast procedures?**

○ No ○ Yes




| Type of procedure: | L | R | B | Date(s) Completed |
|---|---|---|---|---|
| Breast Reduction | ○ | ○ | ○ | _____ |
| Breast Implants | ○ | ○ | ○ | _____ |
| Needle Biopsy | ○ | ○ | ○ | _____ |
| Surgical Biopsy | ○ | ○ | ○ | _____ |
| Lumpectomy | ○ | ○ | ○ | _____ |
| Mastectomy | ○ | ○ | ○ | _____ |
| Breast Reconstruction | ○ | ○ | ○ | _____ |
| Radiation Therapy | ○ | ○ | ○ | _____ |

**Comments:** _____

_____

_____

**Has the Patient ever had breast cancer?**

○ No ○ Yes If yes, age at diagnosis? ☐☐ Age

Which breast? L ○ R ○ B ○

How many **sisters/daughters** with breast cancer? ☐

**Is there a family history of breast cancer?**

If yes, please specify: ○ Mother ○ Sister(s) ○ Other ○ Daughter(s) ☐

○ No ○ Yes ○ Unknown (e.g. adopted)

**Have the Patient's periods stopped permanently?**

○ No ○ Yes ○ Not Sure

If yes or not sure, is she currently taking hormone replacement therapy? ○ No ○ Yes

If yes, how long? _____

4/18/96

49344

# NH Mammography Network
## General Information

A 0 6 4 7 6 0 6

## Instructions:

Please complete this questionnaire using a No.2 pencil or blue or black pen.
All letters and numbers must be written in capital block style **without touching the sides.**

| 0 | 1 | 2 | 3 | 4 | A | B | C | D | E |
|---|---|---|---|---|---|---|---|---|---|

Please shade circles like this: ●

## 1. MAMMOGRAM HISTORY

**Are you having a mammogram today because:** (Choose one)

- ○ **Both you and your health care provider** are concerned about a breast change (lump, pain, etc)?
- ○ **You** are concerned about a breast change?
- ○ **Your health care provider** is concerned about a breast change?
- ○ **Routine Screening Exam** - no breast changes but I or my health care provider wanted a routine mammogram?

**When was your last mammogram?** (Choose one)

- ○ Within the last 12 months
- ○ 1 to 2 years ago
- ○ 3 to 4 years ago
- ○ 5 or more years ago
- ○ Never had a mammogram before

**When did a health care provider last examine your breasts?** (Choose one)

- ○ Within the last 12 months
- ○ 1 to 2 years ago
- ○ 3 to 4 years ago
- ○ 5 or more years ago
- ○ Never

## 2. PERSONAL HISTORY

**What is your date of birth?**

| | | / | | | / | | | | |
|---|---|---|---|---|---|---|---|---|---|
| M | M | | D | D | | Y | Y | Y | Y |

**What is your social security number?**

| | | | - | | | - | | | |
|---|---|---|---|---|---|---|---|---|---|

(To Avoid Duplication of Records)

**What is your racial or ethnic background?** (optional) (Choose one)

- ○ White/Caucasian
- ○ Black/African-American
- ○ Native American (American Indian)
- ○ Hispanic/Latina
- ○ Asian/Pacific Islander
- ○ Other (please specify) _____

**What is your maiden name** (last name only)?

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

**Where were you born?**

○ USA ○ Other _____

**If born in USA, in which state were you born?**

| | |
|---|---|
| | |

State (e.g. NH, VT, MA, ME, etc.)

**What is your current marital status?** (Choose one)

- ○ Single ○ Divorced
- ○ Married ○ Widowed
- ○ Separated

## 2. PERSONAL HISTORY (Contd.)

**What is the highest level of education you have completed?** (Choose one)

- O 8th grade or less
- O Some high school
- O High school graduate
- O Associate's degree or some college/tech school
- O College graduate (4 yrs)
- O Postgraduate

**What is your health insurance coverage?**
(Please shade all that apply)

- O None
- O Private Insurance (Blue Cross, AETNA etc)
- O Medicare
- O Medicaid
- O HMO or PPO (Preferred Provider Organization)
- O CHAMPUS. CHAMPVA or similar
- O Other: _____

**What is your current height?**
(to the nearest inch)

[ ] Feet [ ][ ] Inches

e.g. 5 ft 6½ ins. = 5 0 7

**What is your current weight?**

[ ][ ][ ] Pounds

e.g. 98 lbs. = 0 9 8

**What did you usually weigh (when not pregnant) when you were between 18 and 20 years old?**

[ ][ ][ ] Pounds

## 3. HEALTH HISTORY

**How old were you when you had your first menstrual period?** (Choose one)

- O Under 11
- O 11
- O 12
- O 13
- O 14
- O 15 or older

**Have your Periods stopped permanently?**
O No O Yes

**If Yes, did your Periods stop due to:**
(Choose one)

- O Natural Menopause
- O Surgery (Hysterectomy)
- O Radiation or Chemotherapy
- O Other: _____

**Have you ever had an ovary removed?**
(Choose one)

- O No Ovary Removed
- O Yes, One Ovary Removed
- O Yes, Both Ovaries
- O Yes, but Don't Know if One or Both
- O Don't know

**How old were you at the time of your first full term pregnancy?** (by full term we mean a pregnancy lasting 6 months or more)
(skip if not applicable)

[ ][ ] Age

**How many times have you been pregnant, if ever?** (can be zero)

[ ][ ] Number of Full Term Pregnancies + [ ][ ] Number of Early Pregnancy Losses = [ ][ ] Total Pregnancies



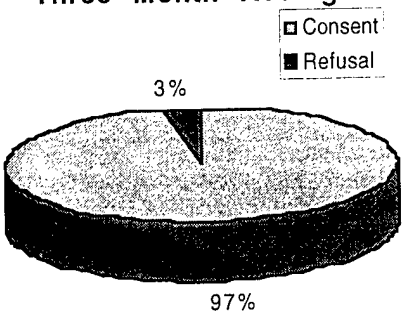
Thank You!

# APPENDIX H – SAMPLE STATUS REPORT FORM (PROCESS MEASURES)

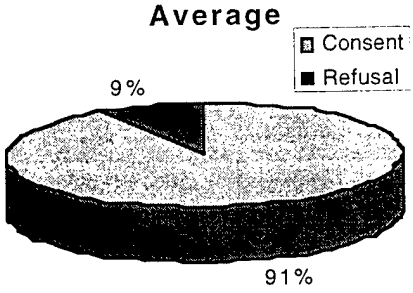
XXX

New Hampshire Mammography Network
NCCC • Evergreen Center, 46 Centerra Parkway, Suite 105, Lebanon, NH 03766-9907
Phone: 603-650-3414 Fax: 603-650-3415

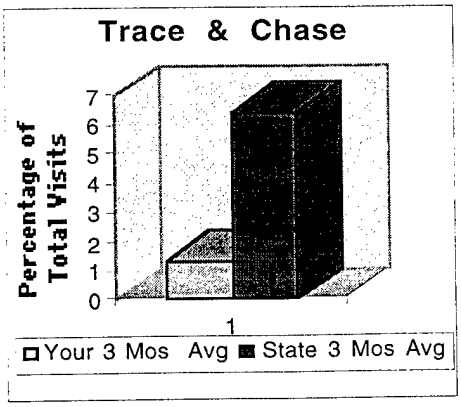
**Your Consent Rate Three Month Average**




**State Aggregate Consent Rate Three Month Average**




**Trace & Chase**




## 3-Month Average At Your Site

The above charts indicate your 3-month average in patient participation (or average number of consents per month) and the completeness of the radiologist forms, compared to the aggregate. Below are your total numbers for the above 3-month period.

Consents/Refusals

| | |
|---|---|
| Consenting Participants | 189 |
| a) Refusals | 10 |
| b) Not Approached | 0 |
| c) Disabled | 0 |

Missing Data

| | |
|---|---|
| No. of Forms with missing data | 4 |
| % of Forms with missing data | 2% |

Missing Forms

We make every effort to match every Patient Intake form we receive with the corresponding signed NHMN consent, however, it is not unusual for mailers to arrive without both forms for each patient. At the end of a 30 day grace period, to allow for the missing form to arrive, we deem those patients to be refusals.

**Definition of terms:**

A consenting participant has signed the NHMN consent form, giving the NHMN permission to enter and track any mammography visits and to link assessment and recommendation with any pathology outcomes.

A refusal indicates that a patient has not signed the NHMN consent form because:

a) having been invited to participate, the patient clearly declined ("Refusals")
or
b) the patient was not asked to complete the survey ("Not Approached")
or
c) the patient could not, through mental or physical disability, understand and sign the form ("Disabled").

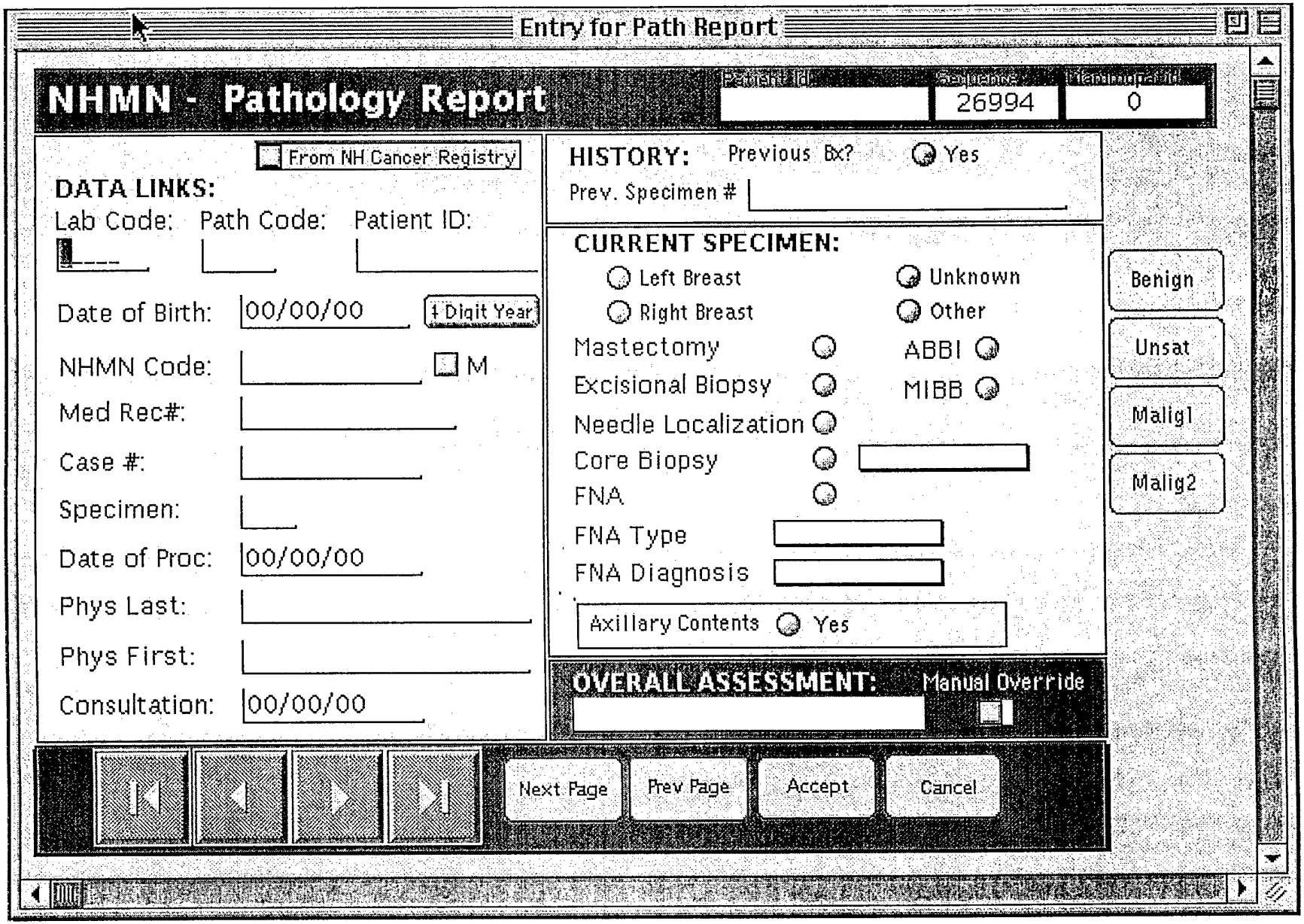
The only identifying data entered for type a), b) or c) is date of birth, date of exam, & zip code.

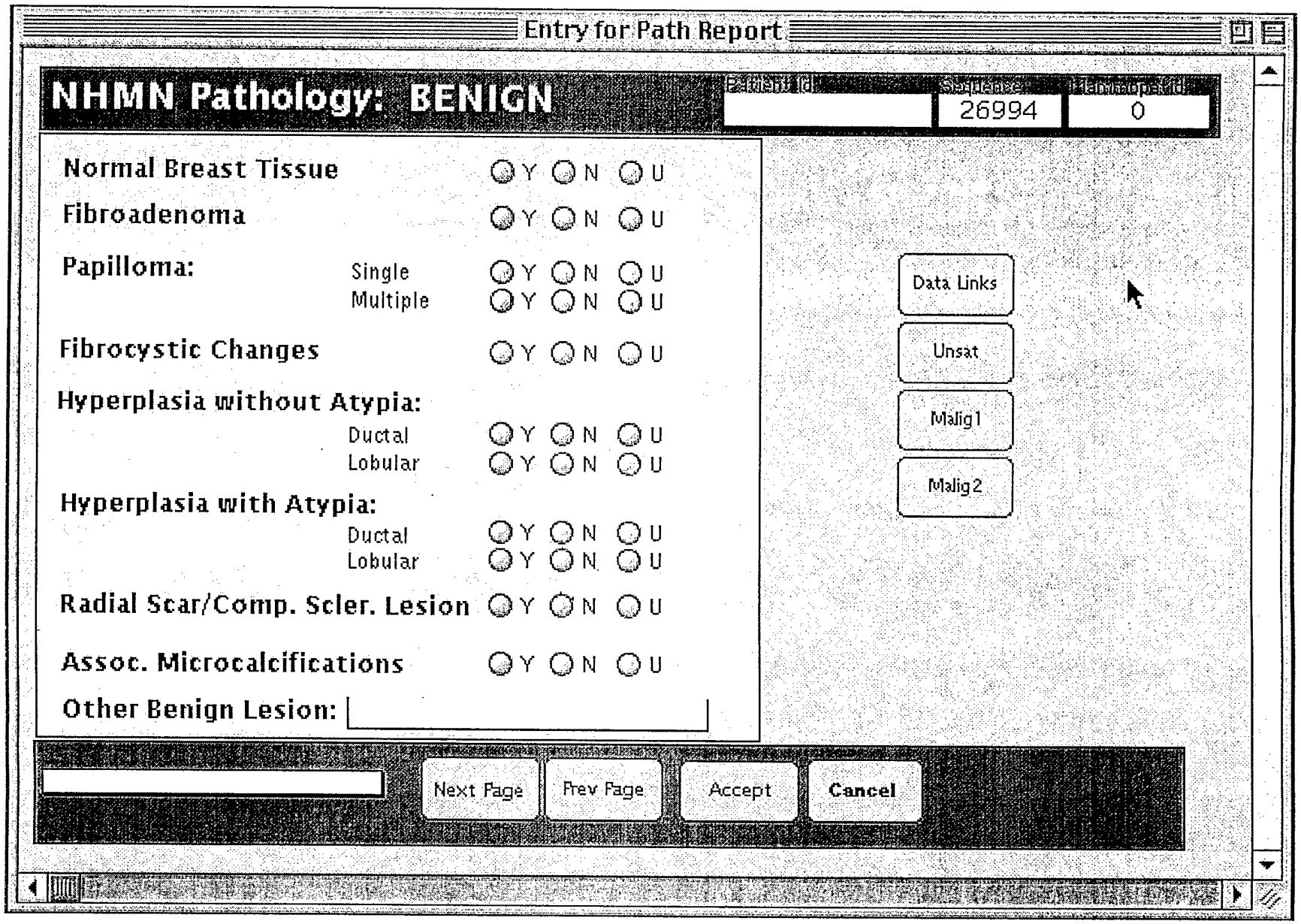
Trace & Chase is an NHMN system for identifying forms with the following essential data items missing:

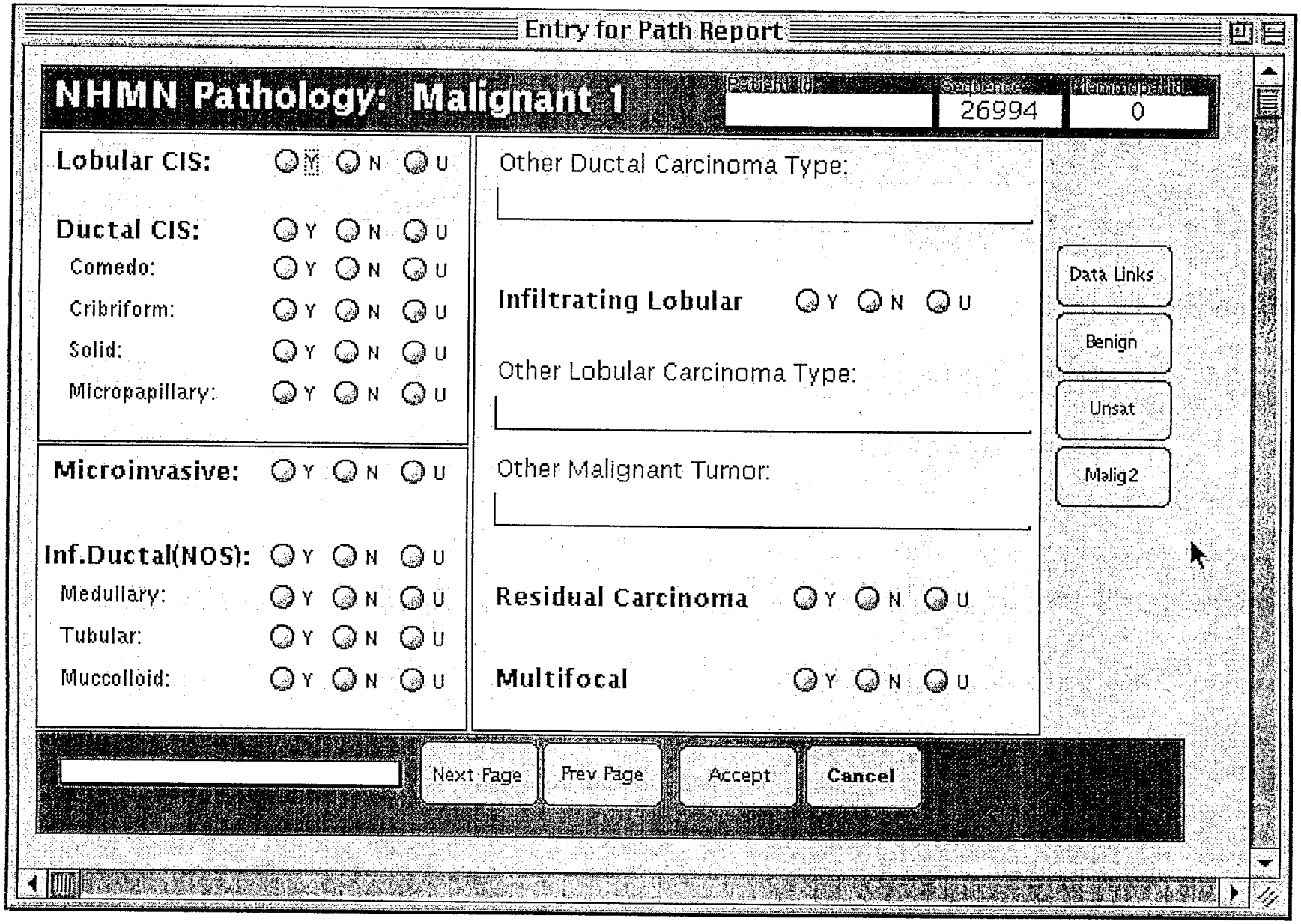
- Rad Initials
- Type of Exam
- Assessment Status
- Recommendation

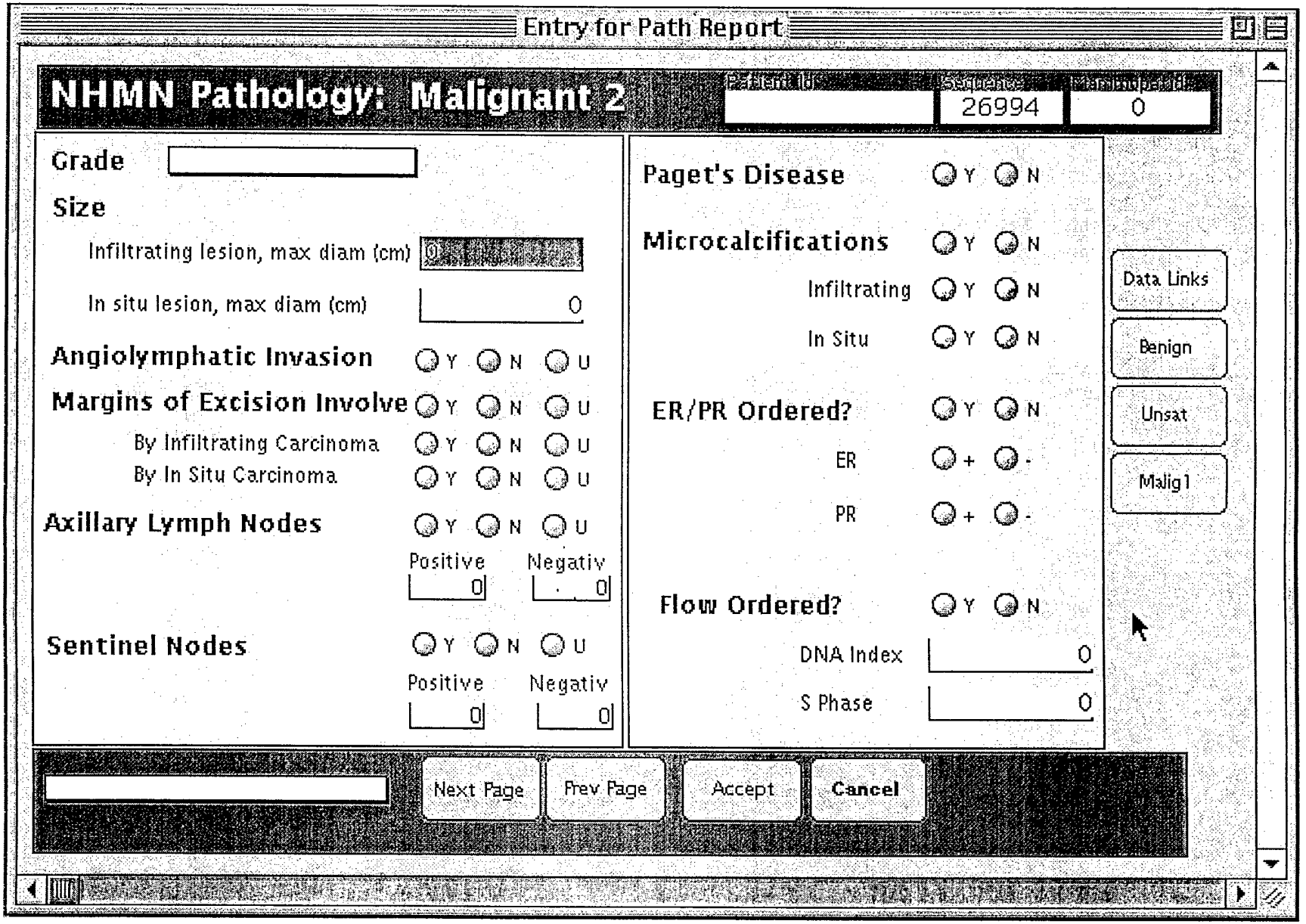
As we identify forms missing one or more of the above items, your site will be given the opportunity to correct your data.

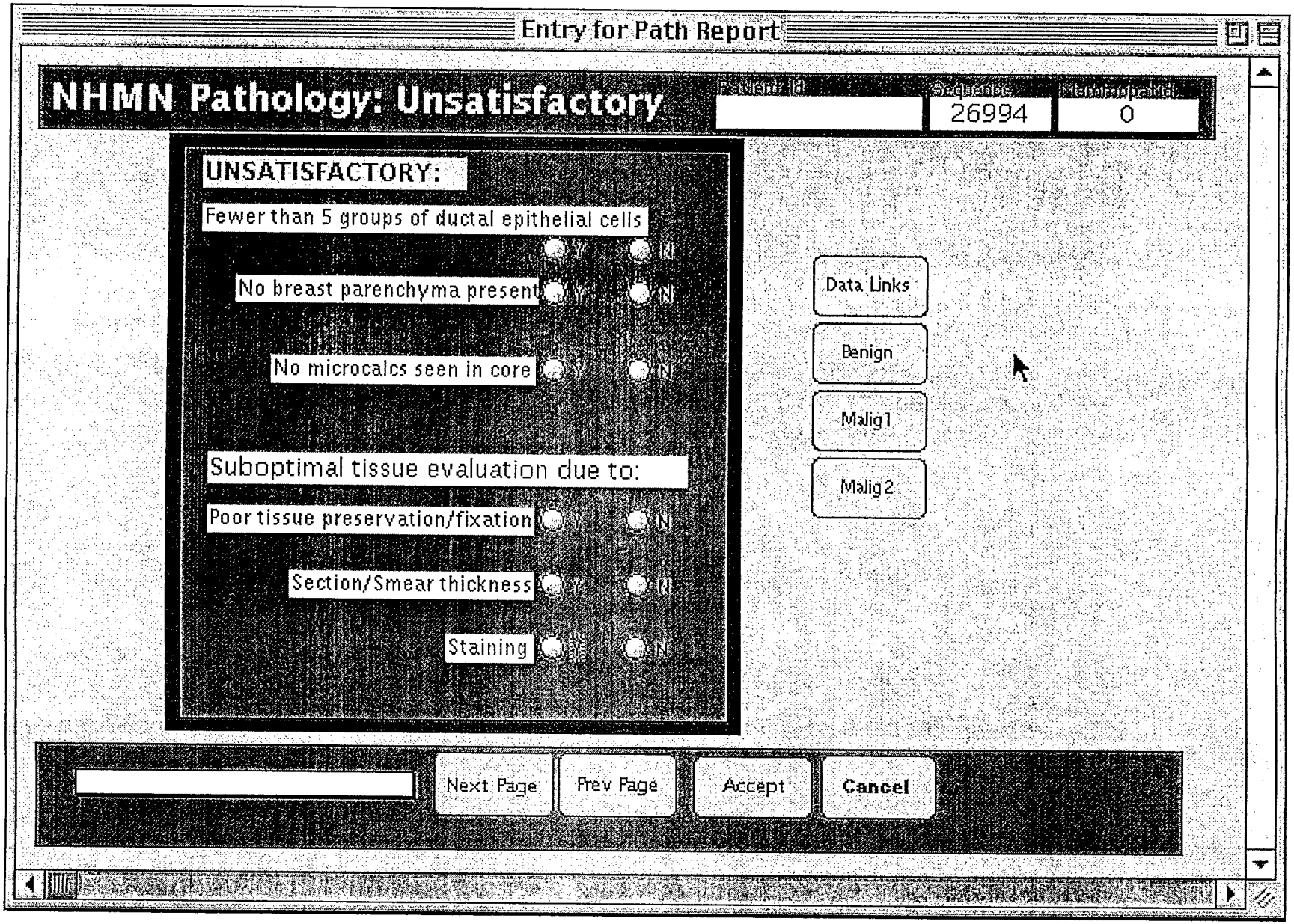
# APPENDIX I- PATHOLOGY INTERPRETATION DATABASE SCREENS


Entry for Path Report
NHMN - Pathology Report
Patient Id
Sequence
26994
Neuropath Id
0
From NH Cancer Registry
DATA LINKS:
Lab Code:
Path Code:
Patient ID:
Date of Birth:
00/00/00
4 Digit Year
NHMN Code:
M
Med Rec#:
Case #:
Specimen:
Date of Proc:
00/00/00
Phys Last:
Phys First:
Consultation:
00/00/00
HISTORY:
Previous Bx?
Yes
Prev. Specimen #
CURRENT SPECIMEN:
Left Breast
Right Breast
Unknown
Other
Mastectomy
Excisional Biopsy
Needle Localization
Core Biopsy
FNA
ABBI
MIBB
FNA Type
FNA Diagnosis
Axillary Contents
Yes
Benign
Unsat
Malig1
Malig2
OVERALL ASSESSMENT:
Manual Override
Next Page
Prev Page
Accept
Cancel


Entry for Path Report
NHMN Pathology: BENIGN
Patient Id
Sequence
26994
Mammogram Id
0
Normal Breast Tissue Y N U
Fibroadenoma Y N U
Papilloma:
Single Y N U
Multiple Y N U
Fibrocystic Changes Y N U
Hyperplasia without Atypia:
Ductal Y N U
Lobular Y N U
Hyperplasia with Atypia:
Ductal Y N U
Lobular Y N U
Radial Scar/Comp. Scler. Lesion Y N U
Assoc. Microcalcifications Y N U
Other Benign Lesion:
Data Links
Unsat
Malig1
Malig2
Next Page
Prev Page
Accept
Cancel


Entry for Path Report
NHMN Pathology: Malignant 1
Patient Id
Sequence
26994
Mammopath
0
Lobular CIS: Y N U
Ductal CIS: Y N U
Comedo: Y N U
Cribriform: Y N U
Solid: Y N U
Micropapillary: Y N U
Microinvasive: Y N U
Inf.Ductal(NOS): Y N U
Medullary: Y N U
Tubular: Y N U
Muccolloid: Y N U
Other Ductal Carcinoma Type:
Infiltrating Lobular Y N U
Other Lobular Carcinoma Type:
Other Malignant Tumor:
Residual Carcinoma Y N U
Multifocal Y N U
Data Links
Benign
Unsat
Malig2
Next Page
Prev Page
Accept
Cancel


Entry for Path Report
NHMN Pathology: Malignant 2
Patient ID
Sequence
26994
Mammogram ID
0
Grade
Size
Infiltrating lesion, max diam (cm)
In situ lesion, max diam (cm)
0
Angiolymphatic Invasion
Y N U
Margins of Excision Involve
Y N U
By Infiltrating Carcinoma
Y N U
By In Situ Carcinoma
Y N U
Axillary Lymph Nodes
Y N U
Positive
Negativ
0
0
Sentinel Nodes
Y N U
Positive
Negativ
0
0
Paget's Disease
Y N
Microcalcifications
Y N
Infiltrating
Y N
In Situ
Y N
ER/PR Ordered?
Y N
ER
+ -
PR
+ -
Flow Ordered?
Y N
DNA Index
0
S Phase
0
Data Links
Benign
Unsat
Malig1
Next Page
Prev Page
Accept
Cancel


Entry for Path Report
NHMN Pathology: Unsatisfactory
Sequence
26994
0
UNSATISFACTORY:
Fewer than 5 groups of ductal epithelial cells
No breast parenchyma present
No microcalcs seen in core
Suboptimal tissue evaluation due to:
Poor tissue preservation/fixation
Section/Smear thickness
Staining
Data Links
Benign
Malig1
Malig2
Next Page
Prev Page
Accept
Cancel

# APPENDIX J – BACKGROUND PUBLICATION ON THE DESIGN AND DEVELOPMENT OF THE NH MAMMOGRAPHY NETWORK

# The New Hampshire Mammography Network: The Development and Design of a Population-Based Registry

Patricia A. Carney[1]
Steven P. Poplack[2]
Wendy A. Wells[3]
Benjamin Littenberg[4]

**OBJECTIVE.** Some authors have proposed a national mammography registry to improve and monitor breast diagnostic practices. However, issues such as confidentiality, accuracy, and direct and indirect costs are practical barriers to implementing such a registry. This paper describes the development and design of a population-based mammography registry in New Hampshire. The project's objectives are to assess the accuracy of mammography by comparing interpretive results with pathology and tumor-registry reports and to improve mammographic performance by reporting findings to facilities, radiologists, and pathologists statewide.

**MATERIALS AND METHODS.** We recruited radiologists and pathologists through professional associations and facilities through site visits. Data used to develop and design the registry were collected during site visits, using structured face-to-face interview methods. Only one site refused to provide site-specific information.

**RESULTS.** Facilities in New Hampshire estimated the annual mammographic volume to be approximately 148,000. We have noted a great deal of variability in mammography practices. Their principal methods for determining screening versus diagnostic mammograms were by patient self-reports (44% of practices), referring physicians' reports (38%), and radiologists' reports (18%). Although 71% of practices have computers, only 16% have radiology information systems or hospital information systems that offer computerized patient-tracking capabilities. More than 90% of New Hampshire radiologists exclusively use freehand dictation for reporting, and although almost 50% codify reports, only 11% use the American College of Radiology lexicon. These data and concerns expressed by radiologists, pathologists, technologists, and administrators helped shape the New Hampshire registry.

**CONCLUSION.** Heterogeneity of radiologic practices poses major challenges for implementing a population-based mammography registry. Issues such as confidentiality, the difficulty of assessing diagnostic acumen, and the time involved in providing data to a registry must be adequately addressed. For the registry to succeed in such diverse settings, researchers, radiologists, pathologists, technologists, and administrative staff must collaborate and cooperate.

Received November 20, 1995; accepted after revision January 18, 1996.

Supported by the United States Department of Defense (DAMD17-94-J-4109) and the National Cancer Institute (CA23108).

[1]Dartmouth Medical School, HB 7920, 1 Medical Center Dr., Lebanon, NH 03756. Address correspondence to P. A. Carney.

[2]Department of Radiology, Dartmouth-Hitchcock Medical Center, HB 7999, 1 Medical Center Dr., Lebanon, NH 03756.

[3]Department of Pathology, Dartmouth-Hitchcock Medical Center, HB 7999, Lebanon, NH 03756.

[4]Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110.

*AJR* 1996;167:367–372

0361–803X/96/1672–367



Development of a national mammography registry was proposed in 1989 as a way to enhance breast-screening effectiveness [1–5]. However, issues of confidentiality, accuracy, direct and indirect costs, and miscommunication erect practical barriers to implementing such a registry [2]. In an attempt to address these concerns, we report the results of an interview survey of radiologists, pathologists, mammography technologists, and administrative staff at mammographic facilities in New Hampshire. The findings from our survey have shaped the design and development of a statewide registry.

New Hampshire has an estimated population of 1,136,000, of whom 160,000 are women 40–74 years old [6]. About 37% of New Hampshire women between 40 and 49 years old report that they have not had a mammogram in the past 2 years, and 50% of women more than 50 years old report no mammogram in the past year [7].

The New Hampshire Mammography Network (NHMN) Project started in October

TABLE 3: Types of Services Provided at Mammography Facilities Participating in the Project ($n$ = 45)

| Services Provided | No. (%) |
|---|---|
| Clinical breast examinations | |
| Routinely provided | 2 (4) |
| Provided to patients with symptoms | 4 (9) |
| Breast sonography | 26 (58) |
| Needle localization | 22 (49) |
| Sonographically guided cyst aspiration | 19 (42) |
| Stereotaxic core biopsies | 5 (11) |

TABLE 4: Methods of Notification Used by Mammographic Facilities in New Hampshire Mammography Network ($n$ = 45)

| Method of Notification | No. (%) |
|---|---|
| Routine mammogram | |
| Notifies patient or primary care provider that mammogram needs to be scheduled | 7 (16) |
| Notifies patient of normal results by mail | 5 (11) |
| Notifies primary care provider of normal results by mail | 45 (100) |
| Abnormal mammogram | |
| Notifies primary care provider by mail | 42 (93) |
| Notifies primary care provider by telephone | 3 (7) |
| Notifies primary care provider by telephone and patient by mail | 5 (11) |
| Biopsy recommendation | |
| Notifies primary care provider by telephone | 26 (58) |
| Notifies primary care provider by mail | 19 (42) |

TABLE 5: Computer Use and Reporting Methods at Mammography Facilities Participating in the Project ($n$ = 45)

| Computer Use and Reporting Methods | No. (%) |
|---|---|
| Type of computer | 32 (71) |
| Macintosh | 5 (16)[a] |
| DOS-based | 27 (84)[a] |
| Radiology information system | 7 (16) |
| Hospital information system | 17 (38) |
| Methods of reporting | |
| Freehand dictation only | 41 (91) |
| Computer generation only | 1 (2) |
| Both | 3 (7) |
| Category system | 22 (49) |
| Site-specific | 18 (82) |
| American College of Radiology | 4 (18) |
| Patient tracking system | |
| Paper-based | 41 (91) |
| Computer-based | 4 (9) |

[a] Percentages are based on the number of facilities that have computers ($n$ = 32).

in such instances is available only when a patient is subsequently biopsied for a palpable abnormality at the same institution, or, in smaller communities, when the facility staff knows the patient.

We investigated notification processes by stratifying reports on the basis of the mammographer's degree of concern (Table 4). Few facilities have systems to remind patients or their primary care providers that routine mammograms are due. Only five facilities (11%) notify patients who are not self-referred of normal results. All facilities routinely contact the requesting physician when a biopsy is recommended, mostly by telephone. The number of radiologists who inform patients of results immediately after the mammogram was not collected.

Almost 40% of New Hampshire hospitals process, section, stain (standard hematoxylin and eosin), and diagnose breast specimens from surgery at their institutions. At 39% of New Hampshire hospitals, breast specimens from surgery are processed, sectioned, and stained at central off-site laboratories, and the slides are returned for diagnosis to the site of surgery. Rural New Hampshire hospitals have breast specimens that are surgically derived at their institutions processed and diagnosed at larger regional institutions. More than 70% of New Hampshire hospitals send fresh breast tumor tissue, when available, to out-of-state commercial laboratories for biochemical analysis of tumor-cell estrogen and progesterone-receptor protein. When diagnostic tissue is limited, paraffin-embedded tissue blocks are sent to the same out-of-state laboratories for immunohistochemical analysis. Almost 30% of New Hampshire hospitals send tissue blocks to a large regional medical center or a state laboratory for immunohistochemical analysis of tumor-cell estrogen and progesterone-receptor protein.

*Staff Concerns at Mammographic Facilities*

The most common concerns about participating in the NHMN Project included confidentiality of data (and attendant medicolegal implications), accuracy of data, and the direct and indirect costs of participation in the project.

Radiologists were universally concerned that participation in the project could expose their practices to damaging legal or public scrutiny. Some feared that plaintiff attorneys might gain access to the registry data and acquire the interpretive results of a particular radiologist in an attempt to show substandard care. Others were worried that collective (statewide) interpretive data might be used to establish standard-of-care norms, which would facilitate malpractice claims. Radiologists were specifically concerned that a lawyer might select data from a particular time range or community to establish a false standard that overestimated the accuracy of mammography. Lastly, some radiologists feared that data might be misused by a particular mammographic facility for marketing purposes. These same concerns were shared by office managers and administrators.

Concerning accuracy of data, radiologists wanted to be certain that data truly reflected their interpretive acumen. Both the accuracy of data entry and the statistical reliability of data were questioned. The issue of statistical reliability was a particular problem because chance alone could profoundly affect a specific radiologist's measures of screening performance if the case load was small.

We heard concerns about the additional work needed for data acquisition and management, and the cost of these services. Technologists worried that collecting patient data for the study would duplicate efforts already performed for site-specific patient-intake forms. Radiologists were concerned that even minimal time spent on each data entry could amount to a significant burden when handling large mammographic volumes. For example, if a radiologist interpretation form took 3 min to complete, then the interpretation of 30 mammograms a day would add 90 min of uncompensated time to each day.

*Registry Design*

Although the design of the registry was fully envisioned at the outset, specifics of data acquisition and implementation were

history, breast surgery history, current breast symptoms, mammography reporting information described with the ACR lexicon, ease of use, affordability, and ability to export data. We have also identified several nonessential features that would be of practical value to the participating mammographic facilities. These features include generation of patient and physician correspondence, the ability to word process dictated reports, the ability to construct reports on the basis of findings present, construction of pathology data fields, and the ability to manage records from multiple mammography sites from a central computer.

We anticipate that many of the high-volume sites may adopt a computerized mammography management system that will encode technologist and radiologist variables and periodically download these data to our centralized database. We hope to offer a system customized to meet the needs of the project as well as the individual sites at a reduced rate. In this customized system, data entry screens on computers would match those on our paper forms.

The concept of offering a computerized mammography management system appeals to personnel at facilities from many perspectives. Such a system allows each facility to act autonomously in the collection and maintenance of interpretive data while capturing more data and decreasing expense for ongoing data acquisition. Accuracy of computerized data entry remains an issue because the project's computer system does not allow double data-entry checks that are often part of a manual registry.

Radiologists were reassured to learn that their recording of interpretations would take less than 1 min and only about 10 sec for 85–95% of interpretations. In addition, we informed facility administrators that both paper-based and computer-assisted data collection options would be available. Many facilities have become particularly interested in computerized systems to limit the handling of multiple paper data collection forms and to facilitate internal interpretive audits of their practices. No matter what the data collection process, however, the project will always lack information on patients who live out of state or refuse to participate.

## Discussion

The NHMN Project shares some of the goals described by Osuch et al. [4] in their proposal for a national mammography database, but our project differs in important ways. We hope to provide an objective assessment of the role of mammography in breast cancer outcomes, and we aspire to improve the accuracy of mammography through a feedback mechanism. One of the major goals of our registry is to create a resource that can be used by health researchers to further our understanding of breast cancer. This objective has not been emphasized in the literature, but we feel it is a critical part of the creation of any mammography database. Though our registry does not assume responsibility for ensuring timely and appropriate patient care, it will monitor long-term outcomes of women receiving mammography.

Many of the criticisms of a national mammography database raised by Taylor and Tocino [5] have been addressed in the development of the NHMN Project, but others present ongoing challenges. Funding has been partly addressed. We were fortunate to receive federal assistance to create the database and to support the central staff. We hope to configure this registry so that once it is functional, it will require minimal funding to maintain. The cost to facilities to participate in this program is difficult to quantify. Clark et al. [8], reporting on the Lee County, FL, mammography registry experience, estimated direct annualized costs of $1.75 for each mammographic report entered, an additional $3936 for each mammography facility, and an additional $1346 for each radiologist. However, no estimate of the indirect costs accrued by the facility and radiology practice was given. The radiology practices we surveyed all appear to operate with only the staff required to perform day-to-day functions; extra time spent on data collection for the project would result in significant expense to the mammographic facility and the radiology practice.

Thus far, participants have willingly given their time without financial compensation. We believe that this support will continue, mostly because the physicians and staff that run mammography facilities have a genuine interest in improving the services they provide. They also aspire to reduce the morbidity and mortality of patients with breast cancer. However, other incentives contribute to their willingness to participate. Many radiologists view participation as a way to satisfy the audit requirements of the Mammography Quality Standards Act of 1992 as administered by the Food and Drug Administration, to gain a more complete understanding of patient-tracking issues, and to measure performance against that of their peers. Also, most mammographers have a strong desire to know how many of their patients with negative mammograms go on to develop breast cancer, a statistic that now is only speculative. We realize that we are in the pilot phase of the project and that enthusiasm may wax and wane as the project progresses, but the fear that the mammography community will be unwilling to participate appears to be unfounded.

The need to standardize mammography and breast pathology reporting is being addressed continually as the project evolves. Our registry follows the ACR lexicon, but it allows radiologists to report on mammograms as they choose. In settings with computerized data acquisition and transcription, this may change, and adoption of the ACR lexicon may become mandatory. We found that most radiologists would be willing to change their reporting methods to comply with the language of the ACR lexicon. Also, we have commitments from all New Hampshire pathologists but one to standardize breast pathology reports.

Taylor and Tocino [5] suggested that a 1-year follow-up period is too soon to detect mammographically occult lesions, which leads to underestimation of the false-negative rate of mammography. We plan to provide statistical analyses that use both 1- and 2-year follow-up periods.

The medicolegal implications of a mammography registry are extensive. We have employed several strategies to protect participants from unnecessary risk, but action at the national level will be required to satisfy all the concerns of participants. We hope that the development of this and other registries will help stimulate federal legislation.

The benefits of a population-based mammography registry include improving the interpretive quality of mammography and improving the follow-up of patients with mammographic abnormalities [4]. We may also further our understanding of breast cancer, including the process of care and the natural history of this disease.

The challenge to implement complex data collection and tracking strategies among mammographic facilities with different organizational structures and staffs who handle high patient volumes is considerable. Meeting quality standards and addressing concerns about confidentiality, accuracy,

# APPENDIX K – SAMPLE FEEDBACK CHARTS (OUTCOME MEASURES)

RESEARCH REPORT - CONFIDENTIAL

Page 1 Mammography Patient Level Tracking (Level 1)
Facility: XXX From: 01/01/98 To: 12/31/98

VOLUMES

| | |
|---|---|
| Total Volume of Mammograms: | 1095 |
| Total Patients with Mammograms: | 1075 |
| Participants (Consenting): | 766 |
| Anonymous (Non Consenting): | 309 |

PATHOLOGY OUTCOMES (Pathology outcomes are only available for Participants)

| | |
|---|---|
| Total # of Participants for whom Pathology Results are Available: | 10 |
| Participants for whom Pathology is available and in whom Cancer was Detected: | 4 |

*** The following reports are attached ***

1. Participants for whom pathology is available 10
2. Patients recommended for biopsy or surgical consult 14
3. Patients with negative mammogram and subsequent cancer developed 1

*** The following reports are available on request ***

4. Patients Recommended for Short Interval Follow Up and/or ACR Category 3 96
5. Patients Recommended for Additional Views to Supplement Current Examination or Breast Ultrasound, and/or ACR Category 0 131

Please note: Data is recorded by patient, not mammogram.

1. All Patients with Pathology Reports:

| Name | DOB | ExamDate | ACR | LAT | ProcDate | LAT | Path Results |
|---|---|---|---|---|---|---|---|
| Doe, J | 03/02/42 | 02/21/97 | 0 | R | 04/24/97 | R | Invasive |
| Doe, J | 08/06/22 | 04/23/97 | 4 | R | 05/28/97 | R | Invasive |
| Doe, J | 06/15/44 | 03/07/97 | 1 | L | 03/26/97 | L | Invasive |
| Doe, J | 01/14/29 | 03/18/97 | 4 | R | 04/23/97 | R | Non-Inva |
| Doe, J | 03/09/36 | 11/25/96 | 0 | L | 09/26/97 | L | Atypical |
| Doe, J | 06/24/47 | 04/07/97 | 2 | L | 04/24/97 | L | Benign |
| Doe, J | 03/30/60 | 03/25/97 | 2 | L | 04/18/97 | L | Benign |
| Doe, J | 07/11/52 | 02/19/97 | 3 | R | 02/25/97 | R | Benign |
| Doe, J | 11/23/70 | 04/11/97 | 1 | R | 04/18/97 | R | Benign |
| Doe, J | 03/24/44 | 11/13/96 | 1 | R | 04/18/97 | R | Benign |

Missing values are shown as periods
refi=referred in; refo=referred out. Please see cover letter.

2. Patients Recommended for Biopsy or Surgical Consultation and/or ACR Category 4 or 5 (Suspicious or Highly Suggestive):

| Name | DOB | ExamDate | ACR | LAT | ProcDate | LAT | Path Results | |
|---|---|---|---|---|---|---|---|---|
| Anonymous, | 07/25/46 | 11/11/96 | 4 | L | . | | | |
| Anonymous, | 09/11/16 | 11/25/96 | 4 | R | . | | | |
| Anonymous, | 06/18/21 | 01/03/97 | 4 | R | . | | | |
| Anonymous, | 07/05/34 | 04/03/97 | 5 | L | . | | | |
| Anonymous, | 09/06/13 | 04/09/97 | 4 | R | . | | | |
| Doe, J | 12/03/56 | 04/18/97 | 4 | R | . | | | refo |
| Doe, J | 08/06/22 | 04/23/97 | 4 | R | 05/28/97 | R | Invasive | |
| Doe, J | 03/01/43 | 02/21/97 | 4 | L | . | | | |
| Doe, J | 04/04/48 | 01/13/97 | 1 | R | . | | | |
| Doe, J | 01/14/29 | 03/18/97 | 4 | R | 04/23/97 | R | Non-Inva | |
| Doe, J | 11/19/53 | 01/20/97 | 4 | L | . | | | refo |
| Doe, J | 05/03/67 | 04/10/97 | 1 | L | . | | | |
| Doe, J | 02/23/52 | 03/18/97 | 4 | L | . | | | refo |
| Doe, J | 09/14/60 | 02/25/97 | 1 | R | . | | | |

Missing values are shown as periods
refi=referred in; refo=referred out. Please see cover letter.

3. Patients with Negative Mammogram and Subsequent Cancer Developed

| Name | DOB | ExamDate | ACR | LAT | ProcDate | LAT | Path Results |
|---|---|---|---|---|---|---|---|
| Doe, J | 06/15/44 | 03/07/97 | 1 | L | 03/26/97 | L | Invasive |

Missing values are shown as periods
refi=referred in; refo=referred out. Please see cover letter.

4. Patients Recommended for Short Interval Follow Up and/or ACR Category 3 (Probably Benign Finding):

| Name | DOB | ExamDate | ACR | LAT | FupDate | FupACR | Path Results |
|---|---|---|---|---|---|---|---|
| Doe, J | 11/05/56 | 04/08/97 | 3 | L | . | . | |
| Doe, J | 06/23/53 | 11/01/96 | 3 | R | (06/05/97) | 3 | |
| Doe, J | 01/01/41 | 11/01/96 | 3 | L | . | . | |
| " | 01/01/41 | 11/01/96 | 3 | R | . | . | |
| Doe, J | 07/18/43 | 11/05/96 | 3 | R | . | . | |
| Doe, J | 06/25/33 | 11/12/96 | 3 | R | . | . | |
| Doe, J | 10/05/43 | 11/11/96 | 3 | R | . | . | |
| Doe, J | 01/27/52 | 01/24/97 | 3 | R | . | . | |
| Doe, J | 05/06/54 | 01/24/97 | 3 | R | . | . | |
| Doe, J | 02/27/44 | 01/23/97 | 3 | L | . | . | |
| " | 02/27/44 | 01/23/97 | 3 | R | . | . | |
| Doe, J | 01/21/22 | 01/20/97 | 3 | L | . | . | |
| Doe, J | 12/26/42 | 01/21/97 | 3 | L | . | . | |
| Doe, J | 06/08/24 | 01/20/97 | 3 | L | . | . | |
| Doe, J | 11/03/45 | 01/20/97 | 3 | R | . | . | |
| Doe, J | 10/12/48 | 12/23/96 | 3 | L | . | . | |
| Doe, J | 12/24/26 | 02/18/97 | 0 | L | . | . | |
| Doe, J | 06/29/37 | 02/17/97 | 3 | L | . | . | |
| Doe, J | 10/18/44 | 02/18/97 | 3 | L | . | . | |
| " | 10/18/44 | 02/18/97 | 3 | R | . | . | |
| Doe, J. | . | 02/25/97 | 3 | R | . | . | |
| Doe, J | 07/06/57 | 01/20/97 | 3 | R | . | . | |
| Doe, J | 10/21/58 | 03/14/97 | 3 | R | . | . | |
| Doe, J | 07/05/34 | 04/03/97 | 3 | R | . | . | |
| Doe, J | 06/07/46 | 03/26/97 | 3 | L | . | . | |
| Doe, J | 10/21/51 | 03/26/97 | 3 | R | . | . | |
| Doe, J | 08/08/46 | 03/06/97 | 3 | L | . | . | |
| Doe, J | 05/17/55 | 04/09/97 | 3 | L | . | . | |
| Doe, J | 09/26/24 | 04/03/97 | 3 | R | . | . | |
| Doe, J | 11/27/25 | 04/07/97 | 3 | R | . | . | |
| Doe, J | 04/09/46 | 04/17/97 | 3 | R | . | . | |
| Doe, J | 07/26/49 | 04/11/97 | 3 | L | . | . | |
| Doe, J | 09/28/47 | 04/10/97 | 3 | L | . | . | |
| Doe, J | 05/07/58 | 04/18/97 | 3 | R | . | . | |
| Doe, J | 09/13/44 | 01/07/97 | 3 | L | . | . | |

Missing values are shown as periods
refi=referred in; refo=referred out. Please see cover letter.

LEVEL 2 RESEARCH REPORT - CONFIDENTIAL

 Mammography Radiologist Final Summary for 1997 05/03/99

Practice xxx

| Rad | Volume | Rad % of Practice | Exam Type | Mammos | Short F/U | BX/Cons Rec. | With Consent | *Results Reported to Date-Mammos with Consent: Cancer Reported: Mammo+ | Cancer Reported: ACR=3 | Cancer Reported: Mammo- | Cancer Reported: Mammo? | Benign Reported: Mammo+ | Benign Reported: ACR=3 | Benign Reported: Mammo- | Benign Reported: Mammo? |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Rxx | 689 | 27.2% | Scrn | 591 | 74(12.5%) | 22( 3.7%) | 521 | 3 | 1 | 0 | 0 | 17 | 65 | 435 | 0 |
| | | | Diag | 98 | 25(25.5%) | 7( 7.1%) | 84 | 1 | 0 | 0 | 0 | 6 | 23 | 54 | 0 |
| Rxx | 629 | 24.8% | Scrn | 547 | 89(16.3%) | 19( 3.5%) | 489 | 3 | 0 | 0 | 0 | 11 | 78 | 397 | 0 |
| | | | Diag | 82 | 12(14.6%) | 16(19.5%) | 76 | 4 | 0 | 0 | 0 | 13 | 12 | 47 | 0 |
| Rxx | 552 | 21.8% | Scrn | 480 | 25( 5.2%) | 22( 4.6%) | 413 | 2 | 0 | 0 | 0 | 18 | 22 | 371 | 0 |
| | | | Diag | 72 | 6( 8.3%) | 5( 6.9%) | 68 | 1 | 0 | 0 | 0 | 4 | 4 | 59 | 0 |
| Rxx | 663 | 26.2% | Scrn | 581 | 73(12.6%) | 20( 3.4%) | 521 | 2 | 0 | 0 | 0 | 18 | 65 | 436 | 0 |
| | | | Diag | 82 | 15(18.3%) | 10(12.2%) | 71 | 3 | 1 | 1 | 0 | 5 | 11 | 50 | 0 |

*This is the final report. 365 days have elapsed for every patient.
Pathology results are linked to initial screening exam if biopsy is within 365 days of exam.
Mammo+=ACR 0,4-5 Mammo-=ACR 1,2 Mammo?=ACR missing.
These data are based on mammograms, not women. Therefore, pathology may be counted more than once.

LEVEL 2 RESEARCH REPORT - CONFIDENTIAL
Mammography Radiologist Final Summary for 1997 05/03/99
Practice xxx

| Rad Volume | Practice Type | Rad % of Exam Mammos | Short F/U | BX/Cons Rec. | With Consent | *Results Reported to Date-Mammos with Consent: Cancer Reported: Mammo+ | ACR=3 | Mammo- | Mammo? | Benign Reported: Mammo+ | ACR=3 | Mammo- | Mammo? |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice Totals for Mammograms | | | | | | | | | | | | | |
| | Scrn | 2199 | 261(11.9%) | 83( 3.8%) | 1944 | 10 | 1 | 0 | 0 | 64 | 230 | 1639 | 0 |
| | Diag | 334 | 58(17.4%) | 38(11.4%) | 299 | 9 | 1 | 1 | 0 | 28 | 50 | 210 | 0 |
| | Total | 2533 | 319(12.6%) | 121( 4.8%) | 2243 | 19 | 2 | 1 | 0 | 92 | 280 | 1849 | 0 |

*This is the final report. 365 days have elapsed for every patient.
Pathology results are linked to initial screening exam if biopsy is within 365 days of exam.
Mammo+=ACR 0,4-5 Mammo-=ACR 1,2 Mammo?=ACR missing.
These data are based on mammograms, not women. Therefore, pathology may be counted more than once.

LEVEL 2 RESEARCH REPORT - CONFIDENTIAL
Mammography Practice Summary for 1997
Practice xxx

05/04/99

| Women | Short F/U | BX/Cons Rec. | Cancer Reported: Mammo+ | Cancer Reported: ACR=3 | Cancer Reported: Mammo- | Cancer Reported: Mammo? | No Cancer Reported: Mammo+ | No Cancer Reported: ACR=3 | No Cancer Reported: Mammo- | No Cancer Reported: Mammo? |
|---|---|---|---|---|---|---|---|---|---|---|
| 1944 | 260( *% ) | 70( 4% ) | 10 | 1 | 0 | 0 | 64 | 230 | 1639 | 0 |

Women Screened

| | ACR=3 is negative | 95% CI | ACR=3 is positive | 95%CI |
|---|---|---|---|---|
| Sensitivity: | 90.9% | * | 100% | * |
| Specificity: | 96.7% | * | 84.8% | * |
| Positive Predictive Value (PPV): | 13.5% | * | 3.6% | * |
| Negative Predictive Value (NPV): | 99.9% | * | 100% | * |

* Estimate not calculated due to insufficent number of values.
Values are based on pathology data received by NHMN.
This page includes one screening mammogram per consenting woman.
ACR=3 is negative if there is no recommendation for followup in less than 9 months.
The first column of statistics counts ACR=3 with less than 9 months followup as negative.
The second column of statistics counts ACR=3 with less than 9 months followup as positive.
No pathology is counted as negative pathology.
CI = Confidence Interval

LEVEL 2 RESEARCH REPORT - CONFIDENTIAL
Mammography Practice Summary for 1997
State Summary

06/04/99

| Women | Short F/U | BX/Cons Rec. | Cancer Reported: Mammo+ | Cancer Reported: ACR=3 | Cancer Reported: Mammo- | Cancer Reported: Mammo? | No Cancer Reported: Mammo+ | No Cancer Reported: ACR=3 | No Cancer Reported: Mammo- | No Cancer Reported: Mammo? |
|---|---|---|---|---|---|---|---|---|---|---|
| 55965 | 2808( 5% ) | 755( 1% ) | 167 | 20 | 50 | 0 | 1510 | 2240 | 51978 | 0 |

Women Screened

| | ACR=3 is negative | 95% CI | ACR=3 is positive | 95%CI |
|---|---|---|---|---|
| Sensitivity: | 70.5% | ( 64.7- 76.3%) | 78.9% | ( 73.7%- 84.1) |
| Specificity: | 97.3% | ( 97.2- 97.4%) | 93.3% | ( 93.1%- 93.5) |
| Positive Predictive Value (PPV): | 10.0% | ( 8.5- 11.4%) | 4.7% | ( 4.1%- 5.4) |
| Negative Predictive Value (NPV): | 99.9% | ( 99.8- 99.9%) | 99.9% | ( 99.9%- 99.9) |

* Estimate not calculated due to insufficent number of values.
Values are based on pathology data received by NHMN.
This page includes one screening mammogram per consenting woman.
ACR=3 is negative if there is no recommendation for followup in less than 9 months.
The first column of statistics counts ACR=3 with less than 9 months followup as negative.
The second column of statistics counts ACR=3 with less than 9 months followup as positive.
No pathology is counted as negative pathology.
CI = Confidence Interval

Sites with no pathology are not included in state totals.

# APPENDIX L – PUBLICATION IN PRESS ON THE PERFORMANCE OF MAMMOGRAPHY IN NEW HAMPSHIRE

# The Performance of Mammography in New Hampshire

Steven P. Poplack, MD*, Anna Tosteson, ScD†**, Meg Grove, MS†, Patricia A. Carney, PhD**

Departments of Radiology*, Medicine †, and Community and Family Medicine** at Dartmouth Medical School, Hanover/Lebanon, NH

**This research was funded by the US Department of Defense (DASD17-94-J-4109).**

## ABSTRACT

**Purpose:** To describe the practice of mammography in a statewide population.

**Materials and Methods:** Mammography data on 47,651 screening and 6,152 diagnostic examinees (11/1/96 - 10/31/97) were linked to 1,572 pathology results. Mammography outcomes were based on BI-RADS assessments and recommendations reported by the interpreting radiologist. The consistency of BI-RADS recommendations was evaluated.

**Results:** Screening mammography had a sensitivity of 72.4% (95%CI: 66.4 - 78.4%), specificity of 97.3% (95%CI: 97.2 -97.4%), and positive predictive value (PPV) of 10.6% (95%CI: 9.0 -12.2%). Diagnostic mammography had higher sensitivity - 78.1% (95%CI: 71.9 -84.3%), lower specificity - 89.3% (95%CI: 88.5 - 90.1%)), and better PPV - 17.1% (95%CI; 14.4 -19.8%). The cancer detection rate of screening was 3.3 per 1000 with a biopsy yield of 22%, while the interval cancer rate was 1.2 per 1000. Nearly 80% of screen detected invasive malignancies were node negative. The recall rate for screening was 8.3%. Ultrasonography was utilized in 3.5% of screening, and 17.5% of diagnostic encounters. BI-RADS recommendations were generally consistent, except for probably benign assessments.

**Conclusions:** The sensitivity of screening mammography in our population- based sample is lower than expected, although other performance indicators are commendable. BI-RADS probably benign assessments are commonly misused.

## INTRODUCTION

The effectiveness of mammography screening in reducing breast cancer mortality, especially in women aged 50 through 69 years old, is well-established (1-9). Numerous studies have evaluated the effectiveness of screening mammography using a variety of outcome measures (10-13), leading to Agency of Health Care Policy and Research (AHCPR) guidelines on the interpretive performance of mammography (14). Most previous studies on interpretive performance involved a limited number of mammography centers with similar characteristics (10-13, 15-18). Few reports have been published on mammography interpretation in diverse community settings (19-22), and to our knowledge no one has described the operating characteristics of both screening and diagnostic mammography or the utilization of breast sonography in a geographically defined largely rural population.

The purpose of this report is to describe key performance measures of screening and diagnostic breast radiography in a geographically defined subject population and to evaluate the use of the American College of Radiology's Breast Imaging Reporting and Data System (BI-RADS) (23) by interpreting radiologists. Our data are derived from a diverse group of mammography facilities, the majority of which are community-based.

## METHODS AND MATERIALS

### Background of the New Hampshire Mammography Network (NHMN)

The design and development of the NHMN is described in detail elsewhere (24, 25). Briefly, the NHMN was founded in October 1994 and began collecting data May 1st, 1996. The NHMN, and all study related procedures, were approved by our committee for the protection of human subjects. All women having mammography in a participating New Hampshire (NH) facility are eligible to enroll in the NHMN. Women participants, radiologists and pathologists

sign written consent to allow for data accrual and analysis by the NHMN. Currently, thirty-seven of the 41 (90%) mammography facilities in NH contribute data to the NHMN. The composition of mammography facilities is diverse and includes hospital (54%) and clinic-based facilities (22%), physician's private offices (20%), free-standing imaging centers (2%), and an academic medical center (2%) (26).

All data contained in the NHMN database are scanned from standardized forms completed by women examinees, mammography technologists and interpreting radiologists. The NHMN does not capture examinations limited to sonography. Examinees provide demographic and some breast cancer risk factor information. Mammography technologists obtain additional risk and clinical information in a face-to-face interview with examinees. During the pilot phase of NHMN development, test-retest reliability studies were conducted on all questions used in data collection for women, including information they provide to technologists during direct interviews. The test-retest results on final data collection forms were greater than 90%.

Radiologists record interpretive data using BI-RADS terminology (23), including: use of breast ultrasound, breast composition, assessment status and recommendation for each breast. We created and distributed to participating radiologists a breast density atlas to assist and standardize coding of radiographic breast density. The atlas displays examples of borderline composition categories (fat vs. scattered density vs. heterogeneously dense vs. extremely dense) and identifies correct density coding for each example. We also conducted quality assurance on interpretive data on 20 randomly selected cases from each facility by comparing data submitted to the NHMN by radiologists to the corresponding clinical text reports. Agreement between NHMN project forms and radiologists' text reports was consistently above 96%.

Participating NH pathology laboratories send clinical pathology reports on all breast specimens, including fine needle aspiration cytology (FNAC), core needle biopsy including the advance breast biopsy instrumentation (ABBI™), excisional biopsy, lumpectomy, and mastectomy, to the NHMN Project office. These are abstracted and entered into a separate pathology database. The most serious pathology outcome is applied when there are multiple pathology results for the same breast, except when a suspicious cytology specimen precedes a benign histology specimen. Linkages between the mammography and pathology databases are performed approximately every six months using a probability-based matching program with demonstrated effectiveness (27).

Study Population

Mammography encounters performed between November 1st, 1996 and October 31, 1997 were eligible for inclusion in these analyses. During this time period, 95 radiologists representing 20 radiology groups interpreted mammography in 36 facilities (87.8%) in NH, and contributed data. We excluded 5,482 women obtaining mammography in six of these mammography facilities because corresponding pathology data were not available for these facilities. We also excluded 805 women who were missing interpretive assessments. Mammography encounter data on 53,803 women were complete and met our inclusion criteria. These were linked with 1,572 benign and malignant pathology results submitted by 82% (14 of 17) of the pathology laboratories in the state of New Hampshire. For 47,651 of these women the initial indication for their exam was screening and for 6,152 women the initial indication was diagnostic.

We defined the nature of a mammography examination based on the presenting indication. We used a hierarchy of the following three independent data sources to identify screening indications: 1) Technologist Form - Examinee reported no current breast concerns

(valid breast concerns were limited to lump, nipple discharge, and skin changes only) and no record in the NHMN database of a prior mammogram of any type within 270 days; 2) Radiologist Form - Type of examination recorded as screening (asymptomatic) mammogram by the interpreting radiologist; 3) Examinee Form - Routine screening exam selected as the indication for mammography. All other examinations not meeting the above criteria were considered diagnostic. The evaluation of a clinical breast concern (pain excluded) and short term (<270 days) follow-up imaging were the primary diagnostic indications. Immediate supplementary imaging (within 45 days of the index screen) was not considered a diagnostic indication, but was linked to the initial screening encounter.

Mammography outcome was based on both the BI-RADS assessment and recommendation reported by the interpreting radiologist. Radiologists recorded assessments and recommendations for each breast, though data were analyzed per woman using the highest assessment category. The BI-RADS assessment category hierarchy was: highly suggestive of malignancy ( category 5) > suspicious abnormality (category 4) > assessment is incomplete (category 0) > probably benign finding (category 3) > benign finding (category 2) > negative (category 1). Mammograms assessed as negative, benign or probably benign with no recommendation for biopsy or surgical consultation were considered negative. Mammograms assessed as highly suggestive of malignancy, suspicious, or incomplete, OR a recommendation for biopsy or surgical consultation irrespective of assessment were considered positive. We analyzed the association of specific recommendations with final assessment categories for each woman. Multiple non-routine recommendations were reported, and may have included a less serious recommendation for the contralateral breast, since recommendations were not analyzed

by laterality. However, recommendations for routine follow-up, non-routine follow-up, and the absence of a recommendation were considered mutually exclusive.

We linked indeterminate screening mammograms (defined as assessment incomplete and/or recommendation for or inclusion of immediate additional evaluation with mammography and/or sonography) with subsequent imaging encounters occurring within 45 days. All linked encounters were considered screening because the initial indication was screening. The outcome of screening mammography reported here reflects the final assessment and recommendation status of associated encounters. An incomplete assessment status implies lack of resolution (within 45 days) of an indeterminate screening mammogram. We limited the time between imaging exams to 45 days after an analysis of 338 initially indeterminate encounters (category 0) revealed that over 98% of women who had supplementary imaging (within 120 days) obtained their exam within 45 days.

We defined recall rate as the proportion of initial screening encounters assessed as incomplete and/or recommending or using additional imaging evaluation to arrive at a final assessment. This did not include definitively abnormal assessments (categories 4 and 5) or probably benign assessments (category 3) rendered solely on the basis of the initial screening encounter.

We defined a positive cancer status as any tissue specimen, including malignant cytology, revealing invasive carcinoma or ductal carcinoma in situ (DCIS). We considered a malignant FNAC outcome to reflect invasive carcinoma. We defined a negative cancer status as a benign result from tissue sampling and/or the absence of malignancy reported within the follow-up interval. Lobular carcinoma in situ (LCIS), atypical epithelial proliferative disorders, and suspicious cytology without correlative histology were considered benign in these analyses.

Statistical Methods

Summary statistics were used to describe patient and examination characteristics for screening and diagnostic mammograms separately. Using the mammography outcome criteria and cancer status definitions above, mammograms were linked with cancer outcomes to identify true positive (TP), true negative (TN), false positive (FP) and false negative (FN) examinations. True positive and false positive status was defined as a positive mammography interpretation with (TP) or without (FP) a cancer diagnosis reported within 365 days. A FP status was incurred irrespective of whether a biopsy was performed. A true negative result was a negative mammography interpretation, including a probably benign assessment , with no report of cancer within the 365day follow-up interval. Similarly a false negative result was defined as a negative mammography interpretation with cancer diagnosed within the subsequent 365day period. Based on these classifications, sensitivity (TP/TP + FN), specificity (TN/FP + TN), positive predictive value (TP/TP + FP), and negative predictive value (TN/FN + TN) were estimated.

Logistic regression, which modeled the odds of a positive mammogram after controlling for cancer status, age <50 vs. ≥ 50 years old, breast density, and history of a prior mammogram (yes vs. no/unknown), was used to account for the influence of varying case-mix on operating characteristics of screening mammography. To facilitate comparisons between sensitivity and specificity in our population with other reports, sensitivity and specificity for women with particular characteristics were estimated using the logistic regression model.

## RESULTS

**Table 1** describes the characteristics of women receiving screening and diagnostic mammography who were included in these analyses. The mean age of the screening population was 54.5 (SD ±11.8) and the median age was 53 years old. Nearly 40% of screened women were

either unsure of their menopausal status or pre-menopausal. The vast majority (92%) of women presenting for screening reported a history of prior mammography.

The recall rate of screening mammography was 8.3%. The final BI-RADS assessments for initially indeterminate screening exams were negative = 2,211 (64.9%), benign = 864 (25.3%), probably benign = 268 (7.9%), suspicious = 62 (1.8%), and highly suggestive of malignancy = 3 (0.1%). Sonography was used or recommended in 3.5% (1,681 of 47,651) of women with a screening indication and 17.5% (1,074 of 6,152) of women with a diagnostic indication. Twenty-three screened women who underwent supplementary imaging evaluation retained an incomplete assessment status, and 516 screened women had no record of additional imaging evaluation. Pathology was available on 130 women with indeterminate index screens, including 42 of the 516 women who had no record of supplementary imaging. There were 28 malignancies reported in the group with additional imaging evaluation and three in the women with no record of supplementary imaging.

**Tables 2a and 2b** list the frequency of final assessments with corresponding recommendations and cancer outcomes for both screening and diagnostic mammography. No recommendation accompanied the assessment in 0.5% of women presenting for screening (n=224 of 47,651) and in 0.8% (n=46 of 6,152) of women presenting for diagnostic mammography. The majority, 90.1% (n=42,925) of screening mammograms were negative (categories 1 or 2) and 98.9% (n=42,440) of negative screens had recommendations for routine follow-up. A smaller proportion, 68.7% (n=4,227), of diagnostic mammograms were considered negative. Approximately 11% (n=472) of negative diagnostic examinations had non-routine recommendations. Suspicious or highly suggestive of malignancy assessments comprised 1.8% (n=842) of screening and 6.5% (n=402) of diagnostic examinations. A recommendation for either

biopsy or surgical consultation accompanied 78.6% (n=602) of suspicious and 92.1% (n=70) of highly suggestive of malignancy screening examinations. This pattern was also seen with diagnostic mammography. Seven percent (n=3,345) of screening mammograms and 21.8% (n=1,341) of diagnostic mammograms were considered probably benign. Less than two thirds of the probably benign assessments (63.1% - screening, and 64.1% - diagnostic) recommended short interval follow-up less than nine months. A small minority of women (1.1% of screening and 3.0% of diagnostic) had incomplete assessments despite the opportunity to resolve this status with supplementary imaging.

**Tables 2a and 2b** also show the frequency of malignancy associated with specific assessment categories. As expected, the frequency of malignancy increases with the severity of the assessment code. Unresolved incomplete screening assessments had a malignancy rate similar to the probably benign category, but were more highly associated with malignancy in women with diagnostic indications. Malignancy was present in less than 2% of the probably benign assessments, which is commensurate with published results (28, 29).

Screening mammography detected malignancy in 3.3 per 1000 women. Diagnostic mammography identified cancer in 21.5 per 1000 patients. Malignancy was diagnosed in 59 women following a negative screening examination, and in 37 women assessed as negative with diagnostic mammography. The interval cancer rate was 1.2/1000 for screening mammography and 6.0/1000 for diagnostic mammography.

**Table 3** outlines sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of screening and diagnostic mammography.

**Table 4** shows estimated sensitivity and specificity of screening mammography according to mammography history, breast density (based on four BI-RADS density categories)

and age (under 50 years versus 50 or older). In our analysis, the odds of a positive screening mammogram increased with increasing breast density. Thus, estimates of sensitivity are higher among women with more dense breasts. In contrast, women with a prior mammogram and women age 50 and older were less likely to have a positive screening exam. Accordingly, mammography was estimated to be less sensitive and more specific among such women. Similar results (not shown) were seen in the diagnostic mammography population.

**Figure 1** illustrates the characteristics of 383 malignancies and details additional staging information on 234 of 319 invasive cancers. The biopsy yield was 22.4% for screening mammography, and 27.5% for diagnostic mammography. Carcinoma in situ accounted for 20.7% of screen-detected malignancy versus 12.1% of cancers identified with diagnostic mammography. Nearly fourteen percent (13.6%) of interval cancers following screening and 21.6% of interval cancers after diagnostic mammography were carcinoma in situ. The mean and median tumor sizes of 88 screen-detected invasive cancers were 16.4mm (SD ±12.1) and 13mm respectively. Almost 80% (70 of 88) of these malignancies did not have axillary lymph node metastases. In contrast the mean and median tumor sizes and axillary node negativity rate of 90 invasive cancers recognized with diagnostic mammography were 22.9mm (SD ± 16.1) and 20mm, and 64.4% respectively. The mean and median tumor sizes and node negativity rate of 36 interval cancers following screening mammography were 17.5mm (SD ± 14.3), 12.5mm and 72.2% respectively. For 20 interval cancers after diagnostic mammography, the mean and median tumor size were 19.6mm (SD ± 15.7) and 16.5mm with a node negativity rate of 80.0%. For interval cancers, the mean time of diagnostic delay (i.e. time from original exam date to pathology date) was 176 days (95%CI: 147-195), with a median of 180 days and range of five to 365 days.

## DISCUSSION

Our data suggest that screening mammography as practiced in a diverse community setting in New Hampshire is considerably less sensitive than the AHCPR published guidelines of 85%. We report mammographic sensitivities ranging from 72.4% to 78.1% and specificities ranging from 89.3% to 97.3%. These sensitivity estimates are lower than most previously reported. However, there are some important methodological differences between our study and other reports (1-13,15-19), which tend to lower sensitivity and raise specificity. We based mammography outcome on the prospective report of the BI-RADS assessment and recommendation encoded by the interpreting radiologist. While BI-RADS is useful for standardization, it is not always used correctly and does not always address complex imaging and clinical circumstances. Unlike most other reports, our mammography results reflect the status of completed imaging evaluations. We classified a probably benign(category 3) assessment as a negative mammography outcome. Almost half, 46%(27 of59), of the interval cancers following screening were assessed as probably benign.

We defined a false negative (interval cancer) result based on the report of a cancer outcome in the 365 days following negative mammography. Our capture of cancer outcomes is enhanced by the inclusion of independent reporting of breast pathology by participating laboratories and may provide a more comprehensive account of positive disease outcomes than other studies that estimate or rely exclusively on tumor registry data (10, 13, 15, 16, 21, 30, 31).

The timing of mammography with respect to clinical breast examination (CBE) may also alter operating characteristics especially sensitivity. Some studies that reported lower interval cancer rates (32, 33) offered CBE at the time of screening mammography, which will decrease

the interval cancer rate due to coincidental detection (by CBE) of mammographically occult cancers.

These methodological differences may help explain our screening sensitivity of 72.4% and corresponding interval cancer rate of 1.2/1000 women. While the sensitivity we report is within the range of sensitivities of 68% - 88% (detection method) noted by Fletcher and colleagues for seven randomized control trials (1), it is lower than reports from single expert centers of 91% to 93% (10, 11). Our sensitivity estimate more closely approximates the rate of 79.9% for linked screening exams noted by the New Mexico Mammography Project (NMMP) (21). Our interval cancer rate of 1.2 per 1000 women is also higher than other published reports (31, 33, 34). Interestingly the tumor sizes and nodal status of our interval cancers were relatively favorable, especially when compared to the staging characteristics of the malignancies identified by diagnostic mammography. We believe this is due to the preponderance of prior mammography (92%) in our screening population. We hypothesize that prior screening may have been effective in extracting larger tumors from the screened population, leaving smaller less detectable cancers available for discovery in the subsequent screen. This may also reflect a clinical decision to perform a biopsy in the settings of a probably benign mammogram or an initially indeterminate exam, which resolved to a negative status (categories 1, 2 or 3) after supplementary imaging was completed.

Our estimates of the sensitivity and specificity of screening mammography based on breast density, history of prior mammography and age identify some interesting and unexpected results. This analysis confirms that prior mammography history is associated with lower sensitivity and higher specificity across all density categories in both age groups. Paradoxically, we note that sensitivity increases with increasing breast density, which contradicts published data

[22]. We speculate that this occurs because denser breasts (i.e. heterogeneously dense and extremely dense) are more likely to be interpreted as positive and engender more intense imaging evaluation then less dense breasts, regardless of the pathology outcome. We found that women with denser breasts had a higher recall rate, greater number of encounters in an imaging series, and higher utilization of breast ultrasound and supplementary mammography.

Our cancer detection rate of 3.3 cancers per 1,000 women screened is comparable to other reports (17, 34-36) given the age distribution and history of prior mammography in our population. One would expect to detect 2-4 cancers per 1,000 women at follow-up or incidence screening and 6-10 cancers per 1,000 women at baseline or prevalence screening (37).

The characteristics of our screen-detected cancers compare favorably with other reports (3, 10, 11, 21, 36, 37, 39). Roughly 21% of our screen detected cancers were carcinoma in situ which is within the range of 19-27% from prior North American reports (10, 11, 21, 36, 37, 39). Mean and median tumor sizes of our invasive cancers were equal or smaller (10, 11, 21, 36, 37, 39). The rate of axillary nodal metastases of 20% for invasive malignancy is also comparable (11, 21, 39), given that studies reporting lower axillary node positivity rates (10, 37) have included carcinoma in situ.

Other measures of screening mammography performance including specificity (97.3%), PPV (10.6%), and recall rate (8.3%) meet AHPCR standards (14). We recognize that these estimates are somewhat inflated by our decision to base mammography outcome on a completed imaging work-up and our definition of recall rate. We defined recall rate according to the guidelines described by Linver et al. (38), which differ from more inclusive abnormality rates, reported by other investigators (10, 19, 21).

In addition to the traditional performance indicators described above we also evaluated the use of BI-RADS by our interpreting physicians. It was reassuring to note that an appropriate recommendation followed the BI-RADS assessment most frequently. However, there were a small but non-trivial number of inappropriate recommendations for all assessments for both screening and diagnostic mammography. Some of these inappropriate recommendations may represent coding errors, indecisiveness resulting in multiple recommendations, or additional non-routine recommendations for the contralateral breast. Some of the discordance reflects the difficulty of applying a rigid coding system to a complex and sometimes ambiguous set of clinical management alternatives. However this also suggests a lack of understanding of BI-RADS by some interpreting radiologists. This was especially evident in the setting of a probably benign assessment, which was associated with a considerable number of routine follow-up recommendations (22% screening and 21% diagnostic) and a higher than expected rate of immediate additional imaging evaluation (14% screening and 11% diagnostic), predominantly ultrasound (11% screening and 10% diagnostic). In these instances the interpreters appear to have misclassified benign and incomplete assessments as probably benign. This underscores the need for training mammographers in the use of BI-RADS, especially as it relates to the appropriate classification and corresponding recommendations of benign, probably benign and incomplete assessments.

While one of the strengths of this report include the collection of standardized data from a diverse group of mammography facilities, the reliance on correct and meticulous coding of data instruments is a potential weakness. Despite the reliability of encoded data noted in our quality assurance analyses, misclassification error remains a concern.

Another concern is the composition of our subject population. Approximately 98% of our study population is Caucasian, which is similar in ethnicity to other population-based mammography databases in the Northeast and the Northwest (40, 41), but differs in ethnic distribution compared to study populations in mammography databases in other regions of the country (21, 35).

## CONCLUSION:

Our data suggest that the sensitivity of screening mammography (72.4%) is lower than generally believed, though other indicators of interpretive performance including cancer detection rate, specificity, PPV (completed imaging work-up), recall rate, and the characteristics of screen detected cancers, satisfy or exceed current standards. Part but not all of the reduction in sensitivity can be explained by the preponderance of prior mammography screening in our population. We also learned that roughly 8% of women presenting for screening mammography will have an indeterminate exam necessitating supplementary imaging evaluation, which will include ultrasonography 23% of the time. Approximately 90% of screening mammograms will be considered negative or definitively benign, 7 % probably benign, 2% suspicious or highly suggestive of malignancy and 1% indeterminate. Appropriate recommendations will follow these assessment categories most of the time, though in the setting of a category 3 (probably benign finding) assessment, recommendations are frequently misapplied. Further education of radiologists in the intended use of the BI-RADS lexicon may help address this problem.

## REFERENCES


1. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the International Workshop on Screening for Breast Cancer. JNCI, 1993,85(20): 1644-1656.
2. Nystrom L, Rutqvist LE, Wall S, Lindgren A, J et al. Breast cancer screening with mammography: Overview of Swedish randomized trials. Lancet, 1993; 341(8851): 973-978.
3. Tabar L, Fagerberg G, Duffy SW. Update of the Swedish two-county program of mammographic screening for breast cancer. Radiologic Clin North Am, 1992; 30:187-210.
4. Roberts MM, Alexander F, Anderson TJ. Edinburgh trail of screening for breast cancer: mortality at seven years. Lancet, 1990; 335:241-246.
5. Frisell J, Klund G, Hellstrom L. Randomized study of mammography screening – preliminary report on mortality in the Stockholm trial. Breast Cancer Res Treat, 1991; 18:49-56.
6. Shapiro S, Venet W, Strax, P. (Eds). Periodic screening for breast cancer. Baltimore: Johns Hopkins University Press, 1988.
7. Andersson I, Aspegren K, Janzon L. Mammographic screening and mortality from breast cancer: the Malmo Mammographic screening trial. BMJ, 1998; 297:943-948.
8. Elwood JM, Cox B, Richardson AK. The effectiveness of breast cancer screening by mammography in younger women. Online Curr Clin Tirals, 1993; 2:Doc. no. 32.
9. Kerlikowske K, Grady D, Rubin SM. Efficacy of screening mammography: a meta-analysis. JAMA, 1995; 273:149-154.
10. Sickles EA, Ominsky SH, Sollitto RA, Galvin HB, Monticciolo DL. Medical Audit of a Rapid-Throughput Mammography Screening Practice: Methodology and Results of 27,114 Examinations. RAD 1990; 175(2): 323-327

11. Bird RE. Low-Cost Screening Mammography: Report on Finances and Review of 21,716 Consecutive Cases. RAD 1989; 171:87-90.

12. Spring DB, Kimbrell-Wilmot K. Evaluating the Success of Mammography at the Local Level: How to Conduct an Audit of Your Practice. RCNA 1987; 25:983-992.

13. Margolin FR, Lagios MD. Development of Mammography and Breast Services in a Community Hospital. RCNA 1987; 25:973-982.

14. Bassett LW, Hendrick, RE, Bassford TL, Butler P, Carter D, Debor M. Quality determinants of mammography. Clinical Practice Guideline, Number 13. Rockville, MD. Agency for Health Care Policy and Research. Public Health Services; 1994; AHCPR Publication Number 95-0632

15. Robertson CL. A Private Breast Imaging Practice: Medical Audit of 25,788Screening and 1,077 Diagnostic Examinations. Radiology 1993; 187:75-79.

16. Wolfe JN, Buck KA, Salane M, Parekh NJ. Xeroradiography of the Breast: Overview of 21,057 Consecutive Cases. Radiology 1987; 165:305-311.

17. Moseson D. Audit of Mammography in a Community Setting. American Journal of Surgery 1992 (May); 163:544-546.

18. Braman DM, Williams HD. ACR Accredited Suburban Mammography Center: Three Year Results. Journal of the Florida Medical Association 1989; 76:1031-1034.

19. Brown ML, Houn F, Sickles EA, Kessler LG. Screening mammography in community practice: positive predictive value of abnormal findings and yield of follow-up diagnostic procedures. AJR 1995; 165:1373-1377.

20. Beam CA, Layde PM, Sullivan DC. Variability in the Interpretation of Screening Mammograms by Radiologists. Arch Intern Med 1996 (Jan); 156:209-213.

21. Rosenberg RD, Lando JF, Hunt WC, Darling RR, Williamson MR, Linver MN, Gilliand FD, Key CR. The New Mexico Mammography Project: Screening Mammography Performance in Albuquerque, NM, 1991 to 1993. Cancer, 1996; 78(8):1731-1739.

22. Rosenberg RD, Hunt WC, Williamson MR et al. Effects of Age, Breast Density, Ethnicity, and Estrogen Replacement Therapy on Screening Mammographic Sensitivity and Cancer Stage at Diagnosis: Review of 183, 134 Screening Mammograms in Albuquerque, New Mexico. Radiology 1998; 209:511-518.

23. Kopans DB, D'Orsi CJ, Adler DED, et al. Breast imaging reporting and data system, 3rd ed. Reston, VA: American College of Radiology, 1998.

24. Carney PA, Poplack SP, Wells et al. Development and Design of a Population-Based Mammography Registry: The New Hampshire Mammography Network American Journal of Roentgenology, 1996;167(August):367-372.

25. Ballard-Barbash R, Taplin SH, Yankaskas BC et al. Breast Cancer Surveillance Consortium: A national mammography screening and outcomes database. American Journal of Roentgenology, 1997; 169:1001-1008.

26. Carney PA, Goodrich ME, O'Mahony D, et al. Mammography in New Hampshire: Characteristics of the Women and the Exams They Receive. (In Review).

27. Automatch, Matchward Tachnologies, Inc. Generalized Record Linkage System (Version 4.2), Kennebunk, Maine, (copyright 1998).

28. Sickles EA. Periodic Mammographic Follow-up of Probably Benign Lesions: Results in 3,184 consecutive Cases. Radiology May 1991;179(2):463-468.

29. Varas X, Leborgne F, Leborgne JH. Nonpalpable, Probably Benign Lesions: Role of Follow-up Mammography. Radiology August 1992;184(2):409-414.

30. Sickles EA. Quality Assurance: How to Audit Your Own Mammography Practice. Radiologic Clinics of North America 1992 (Jan); 30(1):265-275.

31. Burhenne HJ, Burhenne LW, Goldberg F et al. Interval Breast Cancers in the Screening Mammography Program of British Columbia: Analysis and Classification. AJR 1994; 162:1067-1071.

32. Seidman H, Gelb SK, Silverberg E, LaVerda N, Lubera JA. Survival Experience in the Breast Cancer Detection Demonstration Project. CA-A Cancer Journal for Clinicians 1987 (Sept/Oct); 37(5): 258-290.

33. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 1L Breast cancer detection and death rates among women aged 40 to 49 years. Canadian Medical Association Journal 1992; 147(10):1459-1488.

34. Moskowitz M. Interval Cancers and Screening for Breast Cancer in British Columbia – Commentary AJR 1994; 162:1072-1075

35. Kerlikowske K, Grady D, Barclay J, Sickles EA, Eaton A, Ernster V. Positive predictive value of screening mammography by age and family history of breast cancer [see comments]. JAMA. 1993 Nov 24; 270(20): 2444-50.

36. Burhenne LW, Hislop TG, Burhenne HJ. The British Columbia Mammography Screening Program: Evaluation of the First 15 Months. American Journal of Radiology 1992 (Jan); 158:45-49.

37. Linver MN, Paster SB, Rosenberg RD, Key CR, Stidle y CA, King WV. Improvement in Mammography Interpretation Skills in a Community Radiology Practice after Dedicated Teaching Courses: 2-Year Medical Audit of 38,633 Cases. Radiology 1992 (July); 184:39-43.

38. Linver MN, Osuch JR, Brenner JR, Smith RA. The Mammography Audit: A Primer for the Mammography Quality Standards Act (MQSA). American Journal of Radiology 1995(July);165:19-25

39. Morrison AS, Brisson J, Khalid N. Breast Cancer Incidence and Mortality in the Breast Cancer Detection Demonstration Project. Journal of the National Cancer Institute 1988 (Dec.);80(19):1540-1546

40. Geller BM, Worden JK, Ashley JA, Oppenheimer RG, Weaver DL. Multipurpose statewide breast cancer surveillance system: the Vermont experience. J Registry Manage, 1996;23:168-174

41. Thompson RS, Barlow WE, Taplin SH, Grothaus L, Immanuel V, Salazar A, Wagner EH. A population-based case-cohort evaluation of the efficacy of mammographic screening for breast cancer. AJE, 1994; 140(10): 889-901.

Table 1. Characteristics of women receiving mammograms from 11/1/96 through 10/31/97

| Factor | No. of Screening Mammograms N=47,651 n(%) | No. of Diagnostic Mammograms N=6,152 n(%) |
|---|---|---|
| Age (years) | | |
| <40 | 3,230 (6.8) | 1,223 (19.9) |
| 40-49 | 15,468 (32.5) | 2,121 (34.5) |
| 50-59 | 13,753 (28.9) | 1,283 (20.9) |
| 60-69 | 8,880 (18.6) | 853 (13.9) |
| 70-79 | 5,188 (10.9) | 547 (8.9) |
| >79 | 1,130 (2.4) | 120 (2.0) |
| Missing | 2 (<0.01) | 5 (0.1) |
| Ethnicity | | |
| Non-Hispanic White | 31,653 (66.4) | 3,550 (57.7) |
| Other | 630 (1.3) | 72 (1.2) |
| Missing | 15,368 (32.3) | 2,530 (41.1) |
| Prior Mammogram | | |
| None & No date | 3,750 (7.9) | 726 (11.8) |
| > 2 years previously | 13,170 (27.6) | 824 (13.4) |
| 1-2 years previously | 23,252 (48.8) | 1,322 (21.5) |
| <1 year previously | 4,844 (10.2) | 2,842 (46.2) |
| Yes & No date | 2,472 (5.2) | 296 (4.8) |
| Missing | 163 (0.3) | 142 (2.3) |
| Comparison films used | | |
| Yes | 41,487 (87.1) | 4,970 (80.8) |
| No | 4,855 (10.2) | 913 (14.8) |
| Missing | 1,309 (2.8) | 269 (4.4) |
| Menopausal Status (tech form) | | |
| Pre-menopausal | 16,852 (35.4) | 2,753 (44.8) |
| Post-menopausal | 29,136 (61.1) | 2,982 (48.5) |
| Unsure | 1,191 (2.5) | 120 (2.0) |
| Missing | 472 (1.0) | 297 (4.8) |
| Hormone Replacement Therapy | | |
| No | 17,608 (37.0) | 1,815 (29.5) |
| Yes | 11,681 (24.5) | 1,136 (18.5) |
| Missing | 18,362 (38.5) | 3,201 (52.0) |
| Breast Cancer History | | |
| Personal & 1st degree relative | 552 (1.2) | 110 (1.8) |
| Personal | 2,225 (4.7) | 522 (8.5) |
| 1st degree relative | 6,961 (14.6) | 776 (12.6) |
| None | 37,577 (78.9) | 4,504 (73.2) |
| Missing | 336 (0.7) | 240 (3.9) |
| Breast Density | | |
| Fat | 6,580 (13.8) | 581 (9.4) |
| Scattered | 22,125 (46.4) | 2,605 (42.3) |
| Heterogeneously dense | 14,078 (29.5) | 2,001 (32.5) |
| Extremely dense | 4,024 (8.4) | 792 (12.9) |
| Missing | 844 (1.8) | 173 (2.8) |

**Table 2a. Recommendations and cancer outcomes by assessment status for women with screening indications**

| Highest ACR | N (%)[1] | Routine %[2] | Other %[2] | None %[2] | Cancers[3] n(%) |
|---|---|---|---|---|---|
| 1. Negative | 37,995(79.7) | 99.2 | 0.4 | 0.4 | 28(0.1) |
| 2. Benign | 4,930(10.4) | 96.4 | 3.0 | 0.6 | 6(0.1) |
| 3. Prob.Benign | 3,345(7.0) | 22.0 | 77.6 | 0.4 | 27(0.8) |
| 0. Incomplete | 539(1.1) | 6.5 | 89.8 | 3.7 | 3(0.6) |
| 4. Suspicious | 766(1.6) | 3.8 | 95.3 | 0.9 | 104(13.6) |
| 5. Suggestive | 76(0.2) | 4.0 | 94.7 | 1.3 | 46(60.5) |
| Total | 47,651 | 90.8 | 8.8 | 0.5 | 214 |

**Detail for recommendations**

| Highest ACR | N (%)[1] | Routine[2] n (%) | Biopsy[2] n (%) | Surg Consult[2] n (%) | Add'l Views[2] n (%) | Ultra-sound[2] n (%) | Short F/U2 n (%) | CBE[2] n (%) | None[2] n (%) |
|---|---|---|---|---|---|---|---|---|---|
| 1. Negative | 37,995(79.7) | 37,690(99) | 7(<1) | 9(<1) | 23(<1) | 25(<1) | 56(<1) | 46(<1) | 152(<1) |
| 2. Benign | 4,930(10.4) | 4,750(96) | 3(<1) | 1(<1) | 35(1) | 29(1) | 64(1) | 23(<1) | 31(1) |
| 3. Prob.Benign | 3,345(7.0) | 735(22) | 28(1) | 36(1) | 141(4) | 372(11) | 2,110(63) | 92(3) | 13(<1) |
| 0. Incomplete | 539(1.1) | 35(6) | 1(<1) | 5(1) | 214(40) | 277(51) | 34(6) | 10(2) | 20(4) |
| 4. Suspicious | 766(1.6) | 29(4) | 530(69) | 197(26) | 50(7) | 68(9) | 89(12) | 28(4) | 7(1) |
| 5. Suggestive | 76(0.2) | 3(4) | 66(87) | 27(36) | 3(4) | 2(3) | 4(5) | 1(1) | 1(1) |
| Total | 47,651 | 43,242 (90.8) | 635(1) | 275(1) | 466(1) | 773(2) | 2,357(5) | 200(<1) | 224(<1) |

[1]Percentage of women in each assessment category.
[2]Percentage of women with each assessment category that had a specific recommendation
A woman may have more than one non-routine assessment but routine, non-routine, and none are mutually exclusive.
Short follow-up includes only those women with follow-up within 9 months.
When a second mammogram occurs within 45 days, the assessment and recommendations from the second mammogram are used.
[3]Women who developed cancer within one year for a specific assessment category, and % malignant with that assessment.

**Table 2b. Recommendations and cancer outcomes by assessment status for women with diagnostic indications**

| Highest ACR | N (%)[1] | Routine %[2] | Other %[2] | None %[2] | Cancers[3] n(%) |
|---|---|---|---|---|---|
| 1. Negative | 3,026(49.2) | 88.9 | 10.2 | 0.9 | 14(0.5) |
| 2. Benign | 1,201(19.5) | 85.9 | 13.7 | 0.5 | 7(0.6) |
| 3. Prob.Benign | 1,341(21.8) | 21.5 | 78.3 | 0.2 | 19(1.4) |
| 0. Incomplete | 182(3.0) | 4.4 | 92.9 | 2.8 | 8(4.4) |
| 4. Suspicious | 335(5.5) | 1.8 | 97.0 | 1.2 | 64(19.1) |
| 5. Suggestive | 67(1.1) | 0 | 98.5 | 1.5 | 57(85.1) |
| Total | 6,152 | 65.4 | 33.8 | 0.8 | 169 |

**Detail for recommendations**

| Highest ACR | N (%)[1] | Routine[2] n (%) | Biopsy[2] n (%) | Surg Consult[2] n (%) | Add'l Views[2] n (%) | Ultra-sound[2] n (%) | Short F/U[2] n (%) | CBE[2] n (%) | None[2] n (%) |
|---|---|---|---|---|---|---|---|---|---|
| 1. Negative | 3,026(49.2) | 2,691(89) | 14(<1) | 95(3) | 9(<1) | 37(1) | 64(2) | 152(5) | 27(1) |
| 2. Benign | 1,201(19.5) | 1,031(86) | 1(<1) | 15(1) | 6(1) | 13(1) | 108(9) | 32(3) | 6(1) |
| 3.Prob.Benign | 1,341(21.8) | 288(21) | 26(2) | 56(4) | 28(2) | 130(10) | 860(64) | 48(4) | 3(<1) |
| 0. Incomplete | 182(3.0) | 8(4) | 1(1) | 10(5) | 40(22) | 133(73) | 11(6) | 13(7) | 5(3) |
| 4. Suspicious | 335(5.5) | 6(2) | 251(75) | 107(32) | 19(6) | 30(9) | 26(8) | 21(6) | 4(1) |
| 5. Suggestive | 67(1.1) | 0 | 55(82) | 39(58) | 2(3) | 3(4) | 4(6) | 1(1) | 1(1) |
| Total | 6,152 | 4,024(65) | 348(6) | 322(5) | 104(2) | 346(6) | 1,073(17) | 267(4) | 46(1) |

[1]Percentage of women in each assessment category.
[2]Percentage of women with each assessment category that had a specific recommendation
A woman may have more than one non-routine assessment but routine, non-routine, and none are mutually exclusive.
Short follow-up includes only those women with follow-up within 9 months.
When a second mammogram occurs within 45 days, the assessment and recommendations from the second mammogram are used.
[3]Women who developed cancer within one year for a specific assessment category, and % malignant with that assessment.

**Table 3.** Unadjusted Performance Indicators

| | % | 95% CI |
|---|---|---|
| **All women** | | |
| Sensitivity | 74.9 | 70.6-79.3 |
| Specificity | 96.4 | 96.2-96.5 |
| PPV | 12.9 | 11.5-14.3 |
| NPV | 99.8 | 99.8-99.8 |
| **Women with screening mammograms** | | |
| Sensitivity | 72.4 | 66.4-78.4 |
| Specificity | 97.3 | 97.2-97.4 |
| PPV | 10.6 | 9.0-12.2 |
| NPV | 99.9 | 99.9-99.9 |
| **Women with diagnostic mammograms** | | |
| Sensitivity | 78.1 | 71.9-84.3 |
| Specificity | 89.3 | 88.5-90.1 |
| PPV | 17.1 | 14.4-19.8 |
| NPV | 99.3 | 99.1-99.5 |

All women: TP=287, FN=96, TN=51479, FP=1941
Screening: TP=155, FN=59, TN=46135, FP=1302
Diagnostic: TP=132, FN=37, TN=5344, FP=639

Note: Positive mammogram = Assessment categories 0,4, or 5 and/or a recommendation for biopsy or surgical consultation

**Table 4.** Adjusted screening sensitivity and specificity by breast density, age (under 50 and 50 or older), and prior mammogram status.

| Density | Age under 50 | | | | Age 50 or older | | | |
|---|---|---|---|---|---|---|---|---|
| | Prior | | No prior | | Prior | | No Prior | |
| | Sens | Spec | Sens | Spec | Sens | Spec | Sens | Spec |
| Fatty | 66.0 | 98.2 | 74.4 | 97.3 | 63.1 | 98.4 | 72.0 | 97.6 |
| Scattered | 72.8 | 97.5 | 80.1 | 96.3 | 70.2 | 97.8 | 78.0 | 96.7 |
| Heterogen. | 78.0 | 96.7 | 84.2 | 95.1 | 75.8 | 97.1 | 82.5 | 95.8 |
| Ext. Dense | 79.3 | 96.5 | 85.2 | 94.8 | 77.1 | 96.9 | 83.5 | 95.4 |

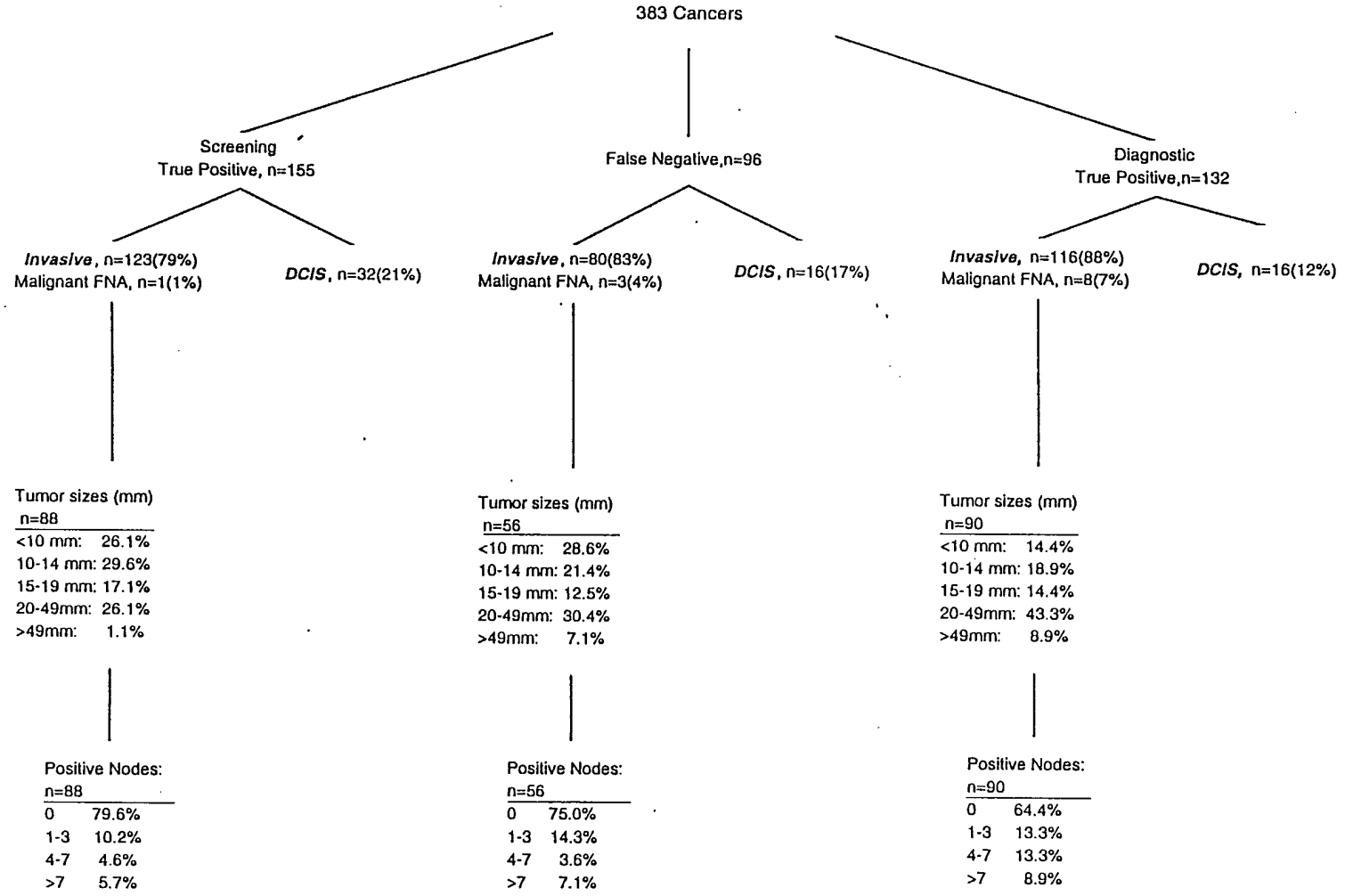



Figure 1. Characteristics of 383 cancers among 53,803 female mammography examinees from 11/1/1996 through 10/31/1997.

# APPENDIX M – PUBLICATION IN PRESS ON THE PERFORMANCE OF MAMMOGRAPHY AMONG WOMEN WITH AND WITHOUT A FIRST DEGREE RELATIVE WITH BREAST CANCER

# PERFORMANCE OF SCREENING MAMMOGRAPHY AMONG WOMEN WITH AND WITHOUT A FIRST-DEGREE RELATIVE WITH BREAST CANCER

Karla Kerlikowske, MD[1,2]

Patricia A. Carney, PhD[3]

Berta Geller, EdD[4]

Margaret T. Mandelson, PhD[5]

Stephen H. Taplin, MD[5]

Kathy Malvin[1]

Virginia Ernster, PhD[1]

Nicole Urban, ScD[6]

Gary Cutter, PhD[7]

Robert Rosenberg, MD[8]

Rachel Ballard-Barbash, MD, MPH[9]



1. Department of Epidemiology and Biostatistics, University of California, San Francisco, CA

2. General Internal Medicine Section, Department of Veterans Affairs, University of California, San Francisco, CA

3. Norris Cotton Cancer Center/Dartmouth-Hitchcock Medical Center/Department of Community and Family Medicine, Dartmouth Medical School, Hanover/Lebanon, NH

4. Office of Health Promotion Research, University of Vermont, College of Medicine, Burlington, VT

5. Center for Health Studies, Group Health Cooperative of Puget Sound, Seattle, WA

6. Cancer Prevention Research Program, Fred Hutchinson Cancer Research Center, Seattle, WA

7. AMC Cancer Research Center, Denver CO

8. University of New Mexico Medical Center, New Mexico Tumor Registry, Albuquerque NM

9. Applied Research Branch, CSRP, DCCPS, National Cancer Institute, Bethesda MD

**Running title**: Family history of breast cancer and screening mammography

**Word count:** 3963

## ABSTRACT

**Background**: Women with a family history of breast cancer are recommended to undergo regular screening mammography beginning at a younger age. Few studies have evaluated the performance of screening mammography among women at increased risk of breast cancer.

**Objective**: To determine the performance of screening mammography in women with a first-degree family history of breast cancer compared to women without of similar age.

**Design**: Cross-sectional.

**Setting**: Seven mammography registries in San Francisco, Seattle, New Hampshire, New Mexico, Vermont, Washington state and Colorado.

**Participants**: 389, 533 women aged 30 to 69 years referred for screening mammography from April, 1985 to November, 1997.

**Measurements**: Breast cancer risk factors, first mammography screening examination captured for a woman by a registry and follow-up of abnormal and normal mammography by linkage to a pathology database or to the Surveillance, Epidemiology, and End Results program or to a state tumor registry to determine occurrence of any invasive cancer or ductal carcinoma *in situ*.

**Results**: The rate of cancer per 1000 examinations increased with age and was higher among women with a family history of breast cancer (3.2 for ages 30-39 [95% CI 1.7, 4.6], 4.7 for ages 40-49 [95% CI 3.8, 5.7], 6.6 for ages 50-59 [95% CI 5.3, 8.0], 9.3 for ages 60-69 [95% CI 7.5, 11.1]; Chi-square for trend P= .001) compared with those without (1.6 for ages 30-39 [95% CI 1.2, 2.0], 2.7 for ages 40-49 [95% CI 2.4, 2.9], 4.6 for ages 50-59 [95% CI 4.1, 5.0], 6.9 for ages 60-69 [95% CI 6.3, 7.5]; Chi-square for trend P= .001). The sensitivity of mammography increased with age among women with a family history of breast cancer (63.2% for ages 30-39 [95% CI 41.5, 84.8], 70.2% for ages 40-49 [95% CI 61.0, 79.5], 81.3% for ages 50-59 [95% CI 73.3, 89.3], 83.8% for ages 60-69 [95% CI 76.8, 90.9]; Chi-square for trend P= .001) and those without (69.5% for ages 30-39 [95% CI 57.7, 81.2], 77.5% for ages 40-49 [95% CI 73.3, 81.8], 80.2% for ages 50-59 [95% CI 76.5, 83.9], 87.7% for ages 60-69 [95% CI 84.8, 90.7]; Chi-

square for trend P= .001) but was similar for each decade of age irrespective of family history status.

**Conclusion**: Having a first-degree relative with a history of breast cancer was associated with cancer detection rates similar to women a decade older without a family history. The sensitivity of screening mammography was primarily influenced by age.

## INTRODUCTION

Guidelines for screening mammography among various organizations recommend that women at high risk of breast cancer undergo regular screening mammography at a younger age than average risk women **Error! Bookmark not defined.**. Few studies have evaluated the performance of screening mammography among young women at increased risk of breast cancer. One group reported the positive predictive value and sensitivity of mammography for women with at least one first-degree relative with a history of breast cancer and found that the positive predictive value was 2 to 3 fold higher **Error! Bookmark not defined.**, yet the sensitivity of mammography was slightly lower compared to women without a family history (3). There are no randomized controlled trials or subgroup analysis of data from existing randomized controlled trials of screening mammography that evaluate the efficacy of screening mammography in women who have a family history of breast cancer.

Understanding whether having a family history of breast cancer influences the performance of mammography may be important in developing screening strategies, especially for younger women where the positive predictive value of mammography is low and the likelihood of associated diagnostic procedures to evaluate an abnormal result is high (2, 4, 5). In order to provide a more stable estimate of the accuracy of mammography among women with a first-degree family history of breast cancer and to compare the accuracy to women without a family history of similar age, we pooled data from seven mammography registries and report the rate of cancer, biopsy yield, and positive predictive value and sensitivity of screening mammography by family history status and decade of age.

## METHODS

### SUBJECTS AND DATA SOURCES

Our study sample included women aged 30 to 69 years who underwent screening mammography between April 1985 and November 1997. Data were pooled from seven mammography registries that participate in the National Cancer Institute Breast Cancer Surveillance Consortium (BCSC) (6). The seven registries are funded by the National Cancer Institute or the Department of Defense and are as follows: 1) San Francisco Mammography Registry (SFMR), 2) Group Health Cooperative (GHC), Seattle, Washington, 3) Fred Hutchinson Cancer Research Center (FHCRC), Seattle, Washington 4) New Mexico, 5) Vermont Mammography Registry (VMR), 6) Colorado Mammography Advocacy Project (CMAP) and 7) New Hampshire Mammography Network (NHMN), Hanover, New Hampshire. The SFMR provided data from April 1985 to December 1993; GHC from January 1986 to December 1993; FHCRC from December 1987 to December 1996; New Mexico from June 1992 to December 1995; VMR from January 1994 to December 1996; CMAP from August 1994 to December 1996, and NHMN from May 1996 to November 1997. Each woman contributed one mammographic examination to the pooled analysis. If a woman had more than one mammographic examination in her respective mammography registry, results from her earliest dated examination were included in the study analyses and results from any subsequent screening examinations were excluded from the study sample. We excluded women with a prior diagnosis of breast cancer or palpable breast mass by history or physical exam. Women with zip codes outside their respective regional Surveillance, Epidemiology, and End Results (SEER) program or state tumor registries' catchment areas were also excluded to minimize incomplete follow-up information.

### MEASUREMENTS

For each woman a self-reported breast cancer risk profile was obtained, as well as a mammographic assessment of two standard screening views per breast. The breast cancer risk profile includes questions about family history of breast cancer in a first-degree relative. Women

are considered to have a family history of breast cancer if they reported having at least one first-degree relative (mother, sister, or daughter) with breast cancer.

Initial screening mammographic assessments were dichotomized into two categories, normal and abnormal. For those mammography registries (SFMR, FHCRC, New Mexico, VMR, NHMN, and CMAP) that used the American College of Radiology Breast Imaging and Reporting Data System (BI-RADS) **Error! Bookmark not defined.** or terminology consistent with BI-RADS to assign mammographic assessment categories, negative (ACR 1) or benign (ACR 2) assessments were classified as normal. Examinations reported with any of the following BI-RADS assessments were categorized as abnormal: 1) probably benign (ACR 3); 2) incomplete, needs additional imaging evaluation (ACR 0), 3) suspicious (ACR 4) and 4) highly suggestive of malignancy (ACR 5). Prior to use of BI-RADS, GHC used three mammographic assessment codes: 1) 'negative', 2) 'indeterminate' and 3) 'positive'. 'Negative' and 'indeterminate' assessments that were recommended for follow-up in one year were classified as normal. 'Indeterminate' assessments that were recommended for six-month follow-up examinations, additional imaging or biopsy and all 'positive' assessments were classified as abnormal.

FOLLOW-UP

Breast biopsies performed as a result of an abnormal mammographic result were identified by contacting the woman's personal physician and/or data linkage with a pathology database and/or data linkage with a radiology database depending on the study site. Breast biopsies included excisional or core biopsies.

Women who underwent screening examinations were linked by computer to a pathology database (VMR, NHMN) and/or to SEER (GHC, SFMR, New Mexico, FHCRC) and/or to a state tumor registry (VMR, NHMN, CMAP) that collects population-based cancer data. Women were linked by using their full name, birth date, address, zip code and social security number when available using probability matching software program (GHC, VMR, NHMN, SFMR; Automatch, Vality Technology, Inc.) or a comparable software program developed for linkage by a

mammography registry (FHCRC, New Mexico, Colorado). Only women who underwent mammography through November, 1997 were eligible for this study to allow at least one year for breast cancers to be detected after normal mammography and to insure that reporting to a tumor registry was complete for all years of the study period. Women were considered to have breast cancer if reports from a breast pathology database, SEER program, or state tumor registry showed any invasive carcinoma or ductal carcinoma *in situ*. Women with lobular carcinoma *in situ* only were excluded. We present results for all breast cancer cases combined and for invasive cancer separately.

## DEFINITIONS AND STATISTICAL ANALYSIS

If breast cancer was diagnosed within 12 months of normal mammography, the normal examination was considered to be a *false negative* examination. If breast cancer was not diagnosed within 12 months of normal mammography, the normal examination was considered to be a *true negative* examination. If breast cancer was diagnosed within 12 months of abnormal mammography, the abnormal examination was considered to be a *true positive* examination. If breast cancer was not diagnosed within 12 months of abnormal mammography, the abnormal examination was considered to be a *false positive* examination. The diagnosis date was the date reported by a SEER program or state tumor registry or the biopsy date recorded in a pathology database.

The positive predictive value of screening mammography was calculated as the percentage of women with abnormal screening examinations who were diagnosed with breast cancer within 12 months of the screening examination. Since the positive predictive value of mammography is influenced by the criteria used to define an examination as 'abnormal', we also report the number of breast cancers detected per 1,000 screening examinations when breast cancer was diagnosed within one year of the screening examination. The cancer yield per breast biopsy was calculated as the percentage of women who underwent breast biopsy who were diagnosed with breast cancer within 12 months of the screening examination. The sensitivity of mammography was calculated

as the number of true positive examinations divided by the number of true positive plus false negative examinations. The specificity of mammography was calculated as the number of true negative examinations divided by the number of false positive plus true negative examinations. The Chi-square test and Fisher's exact test were used for comparison of proportions. The Chi-square test for trend and Chi-square test for homogeneity was used to compare proportions stratified by age. Two-sided P values are reported.

## RESULTS

A total of 389, 533 screening examinations were performed among seven mammography registries, of these 50,834 (13.0%) were performed among women with a family history of breast cancer. Among the five registries that record self-reported prior mammography use, the percentage was similar among women with a family history of breast cancer (81.7%, 28,574/34, 973) compared with women without (80.2%, 170,505/212,729).

### ABNORMAL MAMMOGRAPHY

Among women without a family history of breast cancer, the overall frequency of abnormal examination results was 10.8% (95% CI 10.7, 11.0), ranged from 8.8% to 11.3% across age groups, and was lowest for women aged 30 to 39 years (Table 1). The frequency of abnormal examinations was higher among women with a family history of breast cancer compared with women without (12.7% versus 10.8%; Chi-square $P < .0001$) and these differences were seen for each age group.

### RATE OF CANCER

A total of 1650 breast cancers were identified; 309 (18.7%) were detected among women with a family history of breast cancer (Table 2). The proportion of cancer that was ductal carcinoma *in situ* was similar among women with (22.7%, 95% CI 18.2, 27.8) and without a family history (23.5%, 95% CI 21.3, 25.9). The overall number of cancers detected per screening was 4.2 per 1,000 examinations; 6.1 per 1,000 among women with a family history of breast cancer and 4.0 per 1,000 among women without a family history (Table 2). The number of cancers detected per screening examination increased with age among women with and without a family history of breast cancer (Chi-square for trend; $P= .001$ and $P= .001$, respectively). Women with a family history of breast cancer had 1.5-fold (range 1.3 to 2.0) higher number of cancers detected per screening examination than those without a family history of breast cancer (Chi-square $P < .0001$).

## POSITIVE PREDICTIVE VALUE OF MAMMOGRAPHY

The overall positive predictive value of screening mammography was 3.1% (95% CI 2.9, 3.2) and increased significantly with age from 1.9% for ages 30 to 39 to 6.7% for ages 60 to 69 years for those with a family history of breast cancer and from 1.2% for ages 30 to 39 to 5.6% for ages 60 to 69 years for those without a family history (Chi-square for trend; P= .001 and P= .001, respectively; Table 2). Women with a family history of breast cancer had a slightly higher positive predictive value of mammography compared to those without a family history (3.7% versus 2.9%, Chi-square P= .001).

## BREAST BIOPSIES

The rate of biopsy per screening examination increased with age (Table 3). Women with a family history of breast cancer had a significantly higher rate of biopsy per screening examination compared with women without a family history (16.0 versus 13.1 per 1000 examinations Chi-square P< .0001). The absolute difference in the number of biopsies per examination was smallest among women with and without a family history of breast cancer who were aged 60 to 69 years.

The overall cancer yield per breast biopsy performed was 25.8% (95% CI 24.6, 27.1). The cancer yield per breast biopsy performed increased with age; among women with a family history it was approximately 4 times higher in women aged 60 to 69 years compared to women aged 30 to 39 years (50.6% versus 12.0%; P< .0001); among those without a family history it was 5 times higher in women aged 60 to 69 years compared to women aged 30 to 39 years (40.4% versus 8.4%; P< .0001). Women with a family history of breast cancer had significantly higher yields of breast cancer (invasive breast cancer and ductal carcinoma *in situ* combined) and invasive breast cancer only per breast biopsy performed compared with women without a family history (Chi-square: P = .01 and P = .04, respectively).

## SENSITIVITY OF SCREENING MAMMOGRAPHY

Breast cancers detected by screening mammography among women less than age 50 years were more frequently ductal carcinoma *in situ* (34.8% of cancers) than were those among women age 50 or older (21.4% of cancers) (Chi-square P < .0001). Almost all breast cancers (92.1%) diagnosed within 12 months of mammographic examinations interpreted as normal (false negative examinations) were invasive (Table 4).

The overall sensitivity of screening mammography, allowing 12 months for detection of breast cancer, was 80.9% (95% CI 78.9, 82.8) and increased significantly with age from 63.2% for ages 30 to 39 to 83.8% for ages 60 to 69 years among women with a family history of breast cancer and from 69.5% for ages 30 to 39 to 87.7% for ages 60 to 69 years among women without a family history of breast cancer (Chi-square for trend: P = .006 and P = .001, respectively; Table 4). The sensitivity of screening mammography did not differ significantly among women with and without a family history of breast cancer (77.7% vs. 81.7%, Chi-square P = .1). We also calculated the sensitivity of screening mammography for invasive breast cancer separately. The overall sensitivity for invasive breast cancer was 78.7% (95% CI 76.3, 80.9) and increased significantly with age (Chi-square for trend P= .001; Table 4). The sensitivity for invasive cancer was significantly lower than that for all breast cancers combined for women less than age 50 years (68.6% vs. 74.9%: Chi-square P = .04), but not for those aged 50 and older (83.0% vs. 83.8%; Chi-square P = .6).

The sensitivity of mammography was not associated with the time period that mammography was performed across the various mammography registries (data not shown). The specificity of mammography was lower among women with a family history compared to women without a family history of breast cancer (87.7% versus 89.4%; Chi-square, P < .0001). The specificity of mammography was lower and homogenous among women of all ages with a family history of breast cancer (Chi-square for homogeneity P= .07) and higher and not homogenous among women aged 30 to 69 years without a family history (Chi-square for homogeneity P= .004).

## DISCUSSION

We examined the rate of cancer detection and sensitivity of screening mammography among women with at least one first-degree relative with breast cancer and those without of similar age. The rate of cancer detection per 1,000 screening examinations was 1.3 to 2.0-fold higher among women with a family history of breast cancer compared with women without a family history. In contrast, the sensitivity of mammography was similar irrespective of family history status. Age had a strong effect on sensitivity, being highest among women aged 60 to 69 years (87.0%) and lowest among women aged 30 to 39 years (67.9%). Because this is the largest study to date among women with a family history of breast cancer who have undergone screening mammography, it provides the best estimates for the accuracy of screening mammography in these women.

Other than female gender and older age, having a first-degree relative who has had breast cancer is one of the strongest risk factors for breast cancer **Error! Bookmark not defined.**. The relative risk of breast cancer is 1.5 to 2.4 times higher in women who have a first-degree relative with breast cancer than in women who do not **Error! Bookmark not defined.**. We found the breast cancer detection rate among women with a family history of breast cancer to be similar to women a decade older without a family history of breast cancer. For example, for every 1,000 examinations among women in our study aged 40 to 49 years with a family history of breast cancer, 4.7 cancers were found, which compares with 4.6 per 1,000 examinations among women aged 50 to 59 years without a family history of breast cancer. The higher cancer detection rate we report for women with a family history of breast cancer compared to women without such a history is due to a higher prevalence of breast cancer in these women. Our results are consistent with those of others that show that breast cancer detection rates increase with age **Error! Bookmark not defined.** and are higher among women with a family history of breast cancer compared to those without **Error! Bookmark not defined.**.

The positive predictive value of screening mammography was increased 1.2 to 1.6-fold higher in women with a family history of breast cancer compared to women without a family history. Thus, given an abnormal result, there is only a moderate increase in risk of cancer among

women with a family history. As reported by others (15), we found that for all ages of women the percentage of abnormal results was higher among women with a family history of breast cancer compared with women without. The higher percentage of abnormal results among women with a family history of breast cancer compared to women without may indicate that more breast lesions are actually present among women with a family history or that knowledge of family history alters a radiologist's level of diagnostic suspicion to report a breast lesion. There is some evidence to support the latter explanation. One study has shown that when family history status is known at the time of the mammographic interpretation, radiologists tend to investigate more breast lesions without improving diagnostic accuracy **Error! Bookmark not defined.**. Evaluation of an abnormal mammography result is associated with anxiety up to four months after an abnormal result **Error! Bookmark not defined.** among women with a family history of breast cancer. Determining how availability of information on family history status influences mammographic interpretation may be important to minimize the number of women who may be unnecessarily recalled for diagnostic evaluations and to maximize the positive predictive value of mammography.

As with the percentage of screening examinations interpreted as abnormal, the rate of biopsy was higher among women with a family history of breast cancer compared to those without. However, the absolute difference in the number of breast biopsies per 1000 examinations among women with and without a family history of breast cancer was much smaller compared with the absolute difference in number of abnormal mammography results per 1000 examinations. Given the higher cancer yield per biopsy and only marginally higher positive predictive value of mammography among women with a family history compared to those without suggests that recommending a woman with a family history for a breast biopsy may be a more selective process than recommending her for additional imaging of a mammographic abnormality.

The sensitivity of mammography was primarily influenced by age, not by family history status. Two smaller studies have reported the sensitivity of mammography by age and family history status and showed that sensitivity is slightly lower for women with a family history

compared to women without a family history **Error! Bookmark not defined.**. The sensitivity of first screening mammography has been reported to be higher than for subsequent screening mammography **Error! Bookmark not defined.**. Since women with a family history of breast cancer tend to be somewhat more compliant with routine screening **Error! Bookmark not defined.**, the sensitivity of mammography reported here might be expected to be somewhat lower than that observed in women without a family history. However, given the prevalence of prior screening was high and similar among women with and without a family history of breast cancer, compliance with subsequent screening is unlikely to account for the slightly lower sensitivity of mammography among women with a family history. Alternatively, the slightly lower sensitivity of mammography among women, especially younger women, with a family history of breast cancer compared to those without could be due to a greater proportion of tumors with rapid growth rates that result in higher rates of interval cancers (3).

Our study has several limitations. The accuracy of our data depends on completeness of cancer reporting to the SEER program, state tumor registries, and pathology laboratories at the mammography registries. Also, the registries limit data collection to residents of a defined region. If breast cancer that is detected after a normal mammography examination is not reported to a registry or occurs among women who move out of the data collection region before their breast cancer is diagnosed, false negative examinations may be underestimated, which would result in an overestimation of the sensitivity of mammography. We were not able to calculate screening mammography outcomes by degree of family history, such as number of first-degree relatives with breast cancer or taking into account age at diagnosis of affected relatives, nor did we have information on history of breast cancer among second-degree relatives; thus we could not determine if the performance of screening mammography varies by level of risk. However, 95% of women with a first-degree family history of breast report only one first-degree relative with breast cancer **Error! Bookmark not defined.**. Therefore, it is likely our results are generalizable to the vast majority of women with a family history of breast cancer. We may have underestimated the rate of biopsy per 1,000 examinations since follow-up to determine whether a

biopsy was performed depends on physicians reporting such findings to the registries. However, the rates reported are within the range of those reported in the literature where follow-up has been reported to be 99.6% **Error! Bookmark not defined.**. We report on the performance of first screening examinations captured by a mammography registry which tend to result in higher cancer rates per examination and sensitivity of mammography than subsequent screening examinations (2, 5, 12, 18, 23). As the number of cancers recorded in the BCSC increases, eventually we will be able to report on the performance of mammography among women who have had several examinations within a defined period of time by family history status and age. Lastly, some investigators (4, 5, 24, 25) define an abnormal result as an *initial* BI-RADS assessment of 0, 3, 4 or 5 as defined here while others only consider a *final* assessment of 0, 4, or 5 (12, 23). Consequently, our positive predictive value of mammography may not be generalizable to all medical or mammography practices depending on the definition of an abnormal result used by an individual practice. However, the cancer rate per examination reported here is not influenced by the definition of an abnormal result and is generalizable to all medical and mammography practices.

Our results concern the ability of screening mammography to detect breast cancer among women with and without a family history of breast cancer and do not provide information on the efficacy of screening mammography to reduce breast cancer mortality. However, the lower sensitivity of mammography among younger compared with older women, irrespective of family history status, raises concern about the potential benefit of screening these women. It has been suggested the lower sensitivity of screening mammography observed among younger women may be due to rapid tumor growth rates that result in high rates of interval cancers **Error! Bookmark not defined.**. Given that the identification and treatment of rapid growing tumors (or those that have the potential to differentiate into rapid growing tumors) may have the greatest impact on reducing breast cancer mortality (33), efforts should focus on ways to improve the detection of such tumors.

It has been suggested that younger woman at higher risk of breast cancer, i.e. those with a family history of breast cancer, should discuss with their physicians whether they should begin screening before age 40 **Error! Bookmark not defined.**. In the absence of studies of the efficacy of screening mammography specific to high-risk women aged 30 to 49 years, recommendations for screening such women at a young age have been made on other grounds -- including a high burden of suffering (increased risk of disease and possibly death from breast cancer) **Error! Bookmark not defined.** and a positive predictive value of mammography similar to that of women ages 50 to 69 years **Error! Bookmark not defined.**. Our study results call into question this recommendation given that the sensitivity of mammography did not improve with increased risk, only with increasing age. Similarly, the PPV of mammography primarily increased with age with only a small incremental increase in PPV for high-risk women. Thus, as should be the case for all women, women aged 30 to 49 years with a family of breast cancer should be informed of their individual risk of breast cancer, age-specific chance of an abnormal result, age-specific chance of a false-positive examination, the chance mammography may miss cancer, and the evidence or lack of evidence that screening mammography reduces the risk of death among screened women in their age group.

Table 1: Percentage abnormal result by family history of breast cancer and age

| | AGE A | | | |
|---|---|---|---|---|
| **MEASUREMENTS** | 30 to 39 | 40 to 49 | 50 to 59 | 60 to 69 |
| Number of Exams* | 43,906 | 156,359 | 110,866 | 78,402 |
| Family history† | 6,027 | 19,810 | 13,733 | 11,264 |
| No family history | 37,879 | 136,549 | 97,133 | 67,138 |
| Abnormal % (95% CI) | | | | |
| Family history* | 10.8 | 13.5 | 13.1 | 11.7 |
| | (10.0, 11.5) | (13.1, 14.0) | (12.6, 13.7) | (11.1, 12.3) |
| No family history | 8.8 | 11.3 | 11.0 | 10.8 |
| | (8.5, 9.1) | (11.1, 11.5) | (10.8 11.2) | (10.6, 11.1) |

*First mammography screening examination captured for a woman by a mammography registry.
†Includes women who had at least one first-degree relative (mother, sister, or daughter) with breast cancer.

Table 2: Distribution of breast cancers, rate of breast cancer and positive predictive value of mammography by family history of breast cancer and age

| | AGE A | | | |
|---|---|---|---|---|
| **MEASUREMENTS** | 30 to 39 | 40 to 49 | 50 to 59 | 60 to 69 |
| Breast Cancers | | | | |
| Family history* | | | | |
| Invasive (%) | 13 (68.4) | 70 (74.5) | 74 (81.3) | 82 (78.1) |
| Ductal carcinoma *in situ* † (%) | 6 (31.6) | 24 (25.5) | 17 (18.7) | 23 (21.9) |
| No family history | | | | |
| Invasive % | 40 (67.8) | 259 (71.0) | 342 (77.0) | 385 (81.4) |
| Ductal carcinoma *in situ* † % | 19 (32.2) | 106 (29.0) | 102 (23.0) | 88 (18.6) |
| Breast cancers/1000 exams (95% CI) | | | | |
| Family history*‡ | 3.2 (1.7, 4.6) | 4.7 (3.8, 5.7) | 6.6 (5.3, 8.0) | 9.3 (7.5, 11.1) |
| No family history§ | 1.6 (1.2, 2.0) | 2.7 (2.4, 2.9) | 4.6 (4.1, 5.0) | 6.9 (6.3, 7.5) |
| Positive predictive value mammography % (95% CI) | | | | |
| Family history*‡ | 1.9 (0.8, 2.9) | 2.5 (1.9, 3.0) | 4.1 (3.2, 5.0) | 6.7 (5.3, 8.0) |
| No family history§ | 1.2 (0.9, 1.6) | 1.8 (1.6, 2.0) | 3.3 (3.0, 3.6) | 5.6 (5.1, 6.1) |

*Includes women who had at least one first-degree relative (mother, sister, or daughter) with breast cancer.
†Ductal carcinoma *in situ*

‡Chi-square for trend, P= .001
§ Chi-square for trend, P= .001

Table 3: Rate of biopsy and yield of breast biopsy by family history of breast cancer and age

| | AGE A | | | |
|---|---|---|---|---|
| **MEASUREMENTS** | 30 to 39 | 40 to 49 | 50 to 59 | 60 to 69 |
| Breast biopsies/1000 exams | | | | |
| (95% CI) | | | | |
| Family history* | 16.4 | 15.3 | 17.0 | 15.8 |
| | (13.1, 19.7) | (13.5, 17.1) | (14.7, 19.2) | (13.4, 18.1) |
| No family history | 11.5 | 12.3 | 13.3 | 15.3 |
| | (10.2, 12.7) | (11.7, 13.0) | (12.5, 14.0) | (14.3, 16.3) |
| Breast cancer/biopsy % | | | | |
| (95% CI) | | | | |
| Family history* | | | | |
| All breast cancer | 12.0 | 20.9 | 32.0 | 50.6 |
| | (5.3, 18.6) | (16.2, 25.6) | (25.8, 38.1) | (43.0, 58.2) |
| Invasive cancer only | 6.7 | 14.1 | 24.9 | 39.5 |
| | (1.5, 11.8) | (10.0, 18.1) | (19.1, 30.6) | (32.1, 46.9) |
| No family history | | | | |
| All breast cancer | 8.4 | 16.9 | 28.1 | 40.4 |
| | (5.4, 11.4) | (14.9, 18.8) | (25.5, 30.7) | (37.2, 43.6) |
| Invasive cancer only | 4.9 | 11.0 | 21.5 | 32.3 |
| | (2.6, 7.3) | (9.4, 12.7) | (19.1, 23.9) | (29.3, 35.4) |

*Includes women who had at least one first-degree relative (mother, sister, or daughter) with breast cancer.

Table 4: Sensitivity of screening mammography according to family history of breast cancer and age

| | AGE | | | |
|---|---|---|---|---|
| **MEASUREMENTS** | 30 to 39 | 40 to 49 | 50 to 59 | 60 to 69 |
| Breast Cancers | | | | |
| True positives* | 53 | 349 | 430 | 503 |
| Invasive % | 58.5 | 66.2 | 77.7 | 79.3 |
| Ductal carcinoma *in situ* † % | 41.5 | 33.8 | 22.3 | 20.7 |
| False negatives* | 25 | 110 | 105 | 75 |
| Invasive % | 88.0 | 89.1 | 97.1 | 90.7 |
| Ductal carcinoma *in situ* † % | 12.0 | 10.9 | 2.9 | 9.3 |
| Sensitivity % (95% CI) | | | | |
| Family history‡ | | | | |
| All breast cancer§ | 63.2 | 70.2 | 81.3 | 83.8 |
| | (41.5, 84.8) | (61.0, 79.5) | (73.3, 89.3) | (76.8, 90.9) |
| Invasive cancer only | 53.8 | 62.9 | 77.0 | 81.7 |
| | (26.7, 80.9) | (51.5, 74.2) | (67.4, 86.6) | (73.3, 90.1) |
| No family history | | | | |
| All breast cancer‖ | 69.5 | 77.5 | 80.2 | 87.7 |
| | (57.7, 81.2) | (73.3, 81.8) | (76.5, 83.9) | (84.8, 90.7) |
| Invasive cancer only | 60.0 | 72.2 | 81.0 | 86.2 |
| | (44.8, 75.2) | (66.7, 77.7) | (76.8, 85.2) | (82.8, 89.7) |
| Specificity % (95% CI) | | | | |
| Family history‡ | 89.4 | 86.7 | 87.3 | 89.0 |
| | (88.6, 90.2) | (86.3, 87.2) | (86.8, 87.9) | (88.4, 89.5) |
| No family history | 91.3 | 88.9 | 89.3 | 89.7 |
| | (91.0, 91.6) | (88.7, 89.1) | (89.1, 89.5) | (89.5, 89.9) |

*True positive examination defined as an abnormal examination 12 months prior to a breast cancer diagnosis. False negative examination defined as normal examination 12 months prior to a breast cancer diagnosis
†Ductal carcinoma *in situ*
‡Includes women who had at least one first-degree relative (mother, sister, or daughter) with breast cancer.
§Chi-square for trend, P= .006
‖Chi-square for trend, P= .001

This work was supported by an NCI-funded Breast Cancer Surveillance Consortium co-operative agreement, 1 U01 CA 63740 (to K. Kerlikowske) and K07CA71869 (to M.T. Mandelson) from the Department of Health and Human Services, CA63146 from the National Cancer Institute (to N. Urban) DAMD17-96-1-6288 (to N. Urban) from the Department of Defense.

**Reprint** requests should be addressed to Dr. Karla Kerlikowske, San Francisco Veterans Affairs Medical Center, General Internal Medicine Section, 111A1, 4150 Clement Street, San Francisco CA 94121, Telephone (415) 750-2093, Fax (415) 386-4044, e-mail (**Error! Bookmark not defined.**)

**Authors addresses**

Patricia Carney, PhD
Dept. of Community & Family Medicine
Dartmouth Medical School
1 Medical Center Dr. HB 7925
Lebanon, NH 03756

Berta Geller, PhD
Health Promotion Research
University of Vermont
One South Prospect Street
Burlington, Vt 05401-3444

Meg

Stephen Taplin MD, MPH
Center for Health Studies, Group Health Cooperative
1730 Minor Ave., Suite 1600
Seattle, Wa 98101

Kathy Malvin
U.C.S.F.
74 New Montgomery St., #600
San Francisco, CA 94105

Virginia Ernster, PhD
University of California, San Francisco
400 Parnassus Ave.
Box 0560, MU 420W
San Francsico, CA 94143

Nicole Urban, PhD
Cancer Prevention Research Program
1100 Fairview Ave. N., MP900
P.O. Box 19024
Seattle, Washington 98109-1024

Gary Cutter, Ph.D.
AMC Cancer Research Center
Center for Research Methods and Biometrics
1600 Pierce Street
Lakewood, Colorado 80214

Rob

Rachel Ballard-Barbash, MD, MPH
Associate Director
Applied Research Program
National Cancer Institute
EPN 4005
6130 Executive Blvd- MSC 7344
Bethesda, MD 20892-7344

REFERENCES

**Error! Bookmark not defined.**

# APPENDIX N – ACS SURVEY INSTRUMENT

# Encouraging Mammography Screening in New Hampshire Women


Together with the American Cancer Society, we invite you to participate in a very important study.

With your help, we hope to assist women in obtaining regular mammography, which ultimately could save lives.

*Thank you for your cooperation!*

Patricia A. Carney

Patricia A. Carney, PhD.
Principal Investigator
Norris Cotton Cancer Center
1 Medical Center Dr.
Lebanon, NH 03756

GENERAL INSTRUCTIONS: This booklet contains several sections. Each section represents a new survey which explores your history of obtaining mammograms (x-ray of the breast to identify breast problems) and your feelings about yourself, your risk factors and mammography in general. The answer categories are different for each section. Directions are at the beginning of each new section. If you choose not to answer a particular question, please skip it and move to the next question. The survey should take approximately 25 minutes to complete.

## Section 1. What You Know About Breast Cancer Screening

Directions: For the following questions, please write in or circle the appropriate response.

1.01 When did you last have a mammogram? _____ / _____
Month Year

1.02 Where did you last have a mammogram? _____ / _____
Location State

1.03 How often do women need a mammogram when they are **under 40 years of age?**
every _____ years

1.04 How often do women need a mammogram when they are **40-49 years of age?**
every _____ years

1.05 How often do women need a mammogram when they are **50 years of age or older?**
every _____ years

1.06 At what age do you think the average woman should have her first mammogram?
_____ years old

1.07 If you have discussed mammograms with your doctor, who brought up the issue?

1. My doctor brought up mammograms
2. I brought up mammograms
3. I never discussed mammograms with my doctor

1.08 Medical studies have proven that some groups of women benefit from mammograms. Circle all age groups for which this is true:

1. 18-39 year old women
2. 40-49 year old women
3. 50-74 year old women
4. 75 and older

Study ID #

## Section 2. How You Are Feeling Today

Spielberger SELF-EVALUATION QUESTIONNAIRE©

Directions: A number of statements which people have used to describe themselves are given below. Read each statement then circle the appropriate number to the right of the statement. Do not spend too much time on any one statement but give the answer which best describes how you feel at this moment. There are no right or wrong answers.

| Response Scale | Not At All | Somewhat | Moderately | Very Much So |
|---|---|---|---|---|
| 2.01 I feel calm | 1 | 2 | 3 | 4 |
| 2.02 I feel secure | 1 | 2 | 3 | 4 |
| 2.03 I am tense | 1 | 2 | 3 | 4 |
| 2.04 I feel strained | 1 | 2 | 3 | 4 |
| 2.05 I feel at ease | 1 | 2 | 3 | 4 |
| 2.06 I feel upset | 1 | 2 | 3 | 4 |
| 2.07 I am presently worrying over possible misfortunes | 1 | 2 | 3 | 4 |
| 2.08 I feel satisfied | 1 | 2 | 3 | 4 |
| 2.09 I feel frightened | 1 | 2 | 3 | 4 |
| 2.10 I feel comfortable | 1 | 2 | 3 | 4 |
| 2.11 I feel self-confident | 1 | 2 | 3 | 4 |
| 2.12 I feel nervous | 1 | 2 | 3 | 4 |
| 2.13 I am jittery | 1 | 2 | 3 | 4 |
| 2.14 I feel indecisive | 1 | 2 | 3 | 4 |
| 2.15 I am relaxed | 1 | 2 | 3 | 4 |
| 2.16 I feel content | 1 | 2 | 3 | 4 |
| 2.17 I am worried | 1 | 2 | 3 | 4 |
| 2.18 I feel confused | 1 | 2 | 3 | 4 |
| 2.19 I feel steady | 1 | 2 | 3 | 4 |
| 2.20 I feel pleasant | 1 | 2 | 3 | 4 |

*please turn to next page*

## Section 3. How You Feel in General

Spielberger SELF-EVALUATION QUESTIONNAIRE©

Directions: Read each statement then circle the appropriate number to the right of the statement to indicate how you generally feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which best describes how you generally feel.

| Response Scale | Almost Never | Sometimes | Often | Almost Always |
|---|---|---|---|---|
| 3.01 I feel pleasant | 1 | 2 | 3 | 4 |
| 3.02 I feel nervous and restless | 1 | 2 | 3 | 4 |
| 3.03 I feel satisfied with myself | 1 | 2 | 3 | 4 |
| 3.04 I wish I could be as happy as others seem to be | 1 | 2 | 3 | 4 |
| 3.05 I feel like a failure | 1 | 2 | 3 | 4 |
| 3.06 I feel rested | 1 | 2 | 3 | 4 |
| 3.07 I am "calm, cool and collected" | 1 | 2 | 3 | 4 |
| 3.08 I feel that difficulties are piling up so that I cannot overcome them | 1 | 2 | 3 | 4 |
| 3.09 I worry too much over something that really doesn't matter | 1 | 2 | 3 | 4 |
| 3.10 I am happy | 1 | 2 | 3 | 4 |
| 3.11 I have disturbing thoughts | 1 | 2 | 3 | 4 |
| 3.12 I lack self-confidence | 1 | 2 | 3 | 4 |
| 3.13 I feel secure | 1 | 2 | 3 | 4 |
| 3.14 I make decisions easily | 1 | 2 | 3 | 4 |
| 3.15 I feel inadequate | 1 | 2 | 3 | 4 |
| 3.16 I am content | 1 | 2 | 3 | 4 |
| 3.17 Some unimportant thought runs through my mind and bothers me | 1 | 2 | 3 | 4 |
| 3.18 I take disappointments so keenly that I can't put them out of my mind | 1 | 2 | 3 | 4 |
| 3.19 I am a steady person | 1 | 2 | 3 | 4 |
| 3.20 I get in a state of tension or turmoil as I think over my recent concerns & interests | 1 | 2 | 3 | 4 |

## Section 4. How You Feel About Mammography

Directions: Below is a list of comments women have made about mammography. Please read each item and indicate how strongly you agree by circling the response that best fits your situation.

| Response Scale: | Strongly Agree | Moderately Agree | Neutral | Moderately Disagree | Strongly Disagree |
|---|---|---|---|---|---|
| 4.01 When I get a recommended mammogram, I feel good about myself | 1 | 2 | 3 | 4 | 5 |
| 4.02 When I get a mammogram I don't worry as much about breast cancer | 1 | 2 | 3 | 4 | 5 |
| 4.03 Having a mammogram will help me find breast lumps early | 1 | 2 | 3 | 4 | 5 |
| 4.04 Having a mammogram will decrease my chances of dying from breast cancer | 1 | 2 | 3 | 4 | 5 |
| 4.05 Having a mammogram will decrease my chances of requiring radical or disfiguring surgery if breast cancer occurs | 1 | 2 | 3 | 4 | 5 |
| 4.06 Having a mammogram will help me find a lump before it can be felt by myself or a health professional | 1 | 2 | 3 | 4 | 5 |
| 4.07 Having a routine mammogram would make me worry about breast cancer | 1 | 2 | 3 | 4 | 5 |
| 4.08 Having a mammogram would be embarrassing | 1 | 2 | 3 | 4 | 5 |
| 4.09 Having a mammogram would take too much time | 1 | 2 | 3 | 4 | 5 |
| 4.10 Having a mammogram would be painful | 1 | 2 | 3 | 4 | 5 |
| 4.11 Having a mammogram would cost too much money | 1 | 2 | 3 | 4 | 5 |

*please turn to next page*

## Section 5. How You Feel About Stressful Life Events

**Directions: Below is a list of comments made by people after stressful life events. Please read each item, then circle the number that most accurately describes how frequently these comments were true for you during the past 7 days. If they did not occur during that time, please circle the number 1 for "not at all." "It" refers to the stressful event.**

| Response Scale: | Not at All | Rarely | Sometimes | Often |
|---|---|---|---|---|
| 5.01 I thought about it when I didn't mean to | 1 | 2 | 3 | 4 |
| 5.02 I avoided letting myself get upset when I thought about it or was reminded of it | 1 | 2 | 3 | 4 |
| 5.03 I tried to remove it from memory | 1 | 2 | 3 | 4 |
| 5.04 I had trouble falling asleep or staying asleep because of pictures or thoughts about it that came into my mind | 1 | 2 | 3 | 4 |
| 5.05 I had waves of strong feelings about it | 1 | 2 | 3 | 4 |
| 5.06 I had dreams about it | 1 | 2 | 3 | 4 |
| 5.07 I stayed away from reminders of it | 1 | 2 | 3 | 4 |
| 5.08 I felt as if it hadn't happened or it wasn't real | 1 | 2 | 3 | 4 |
| 5.09 I tried not to talk about it | 1 | 2 | 3 | 4 |
| 5.10 Pictures about it popped into my mind | 1 | 2 | 3 | 4 |
| 5.11 Other things kept making me think about it | 1 | 2 | 3 | 4 |
| 5.12 I was aware that I still had a lot of feelings about it, but didn't deal with them | 1 | 2 | 3 | 4 |
| 5.13 I tried not to think about it | 1 | 2 | 3 | 4 |
| 5.14 Any reminder brought back feelings about it | 1 | 2 | 3 | 4 |
| 5.15 My feelings about it were kind of numb | 1 | 2 | 3 | 4 |

## Section 6. How You Feel About Cancer Screening

**Directions: We would like to understand your attitudes about cancer in general. Please circle the appropriate response in each of the following statements.**

| Response Scale: | Strongly Agree | Somewhat Agree | Somewhat Disagree | Strongly Disagree |
|---|---|---|---|---|
| 6.01 Even though it's a good idea, I find that getting an examination for cancer scares me. | 1 | 2 | 3 | 4 |

*Section 6 continued...*

| Response Scale: | Strongly Agree | Somewhat Agree | Somewhat Disagree | Strongly Disagree |
|---|---|---|---|---|
| 6.02 Even though it's a good idea, I find that having my breasts examined is embarrassing | 1 | 2 | 3 | 4 |
| 6.03 When I see a news story about cancer I usually skip it without reading it | 1 | 2 | 3 | 4 |
| 6.04 The word cancer scares me | 1 | 2 | 3 | 4 |
| 6.05 If I got cancer, I'd rather not know about it | 1 | 2 | 3 | 4 |
| 6.06 If doctors find cancer, there's nothing they can do anyway | 1 | 2 | 3 | 4 |
| 6.07 Since no one knows what causes cancer, there's really nothing that can be done about it | 1 | 2 | 3 | 4 |
| 6.08 Getting cancer is a death sentence for most people | 1 | 2 | 3 | 4 |
| 6.09 I know they say finding cancer early is a good idea, but I'd rather not have it checked | 1 | 2 | 3 | 4 |
| 6.10 Once a person develops cancer, it is usually too late to do anything about it | 1 | 2 | 3 | 4 |
| 6.11 I think they will find a cure for cancer | 1 | 2 | 3 | 4 |
| 6.12 If most people got health checkups regularly, there would be fewer deaths from cancer | 1 | 2 | 3 | 4 |
| 6.13 Cancer doesn't cause death as often as most people think | 1 | 2 | 3 | 4 |
| 6.14 I think if I got cancer, I could make a pretty good adjustment | 1 | 2 | 3 | 4 |
| 6.15 If I had cancer, I would still enjoy being friends | 1 | 2 | 3 | 4 |
| 6.16 I would still feel life is worth living even though I have cancer | 1 | 2 | 3 | 4 |
| 6.17 If I got cancer I would feel okay around people who didn't have it | 1 | 2 | 3 | 4 |
| 6.18 If I had cancer, being treated with drugs and/or radiation would be worth all the side effects because it might save my life | 1 | 2 | 3 | 4 |

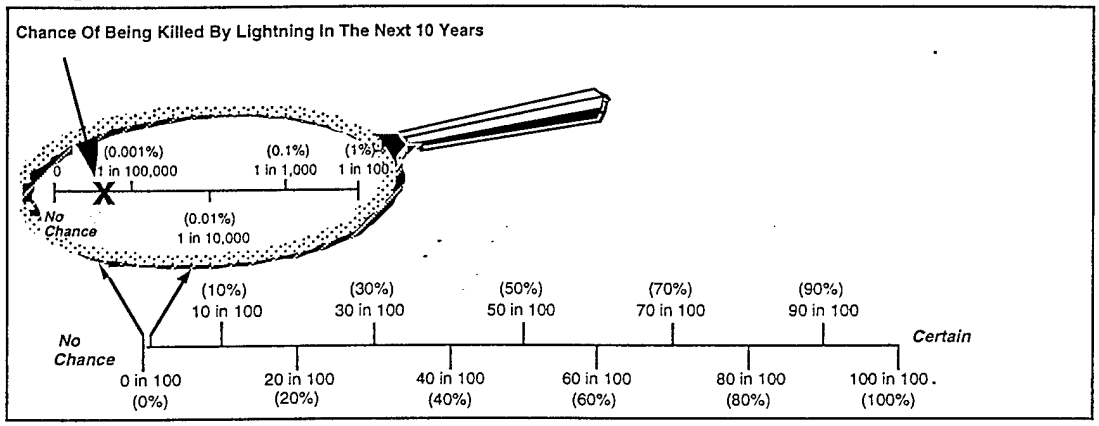
*please turn to next page*

## Section 7. What You Think About Cancer Risk

Directions: On the next few pages you will find questions about how likely it is that various things will happen. We will ask you to put your answers on scales like the ones that follow. The scale is a line which goes from "no chance" (0%) to "certain" (100%). It has a magnifying glass for the smallest chances.

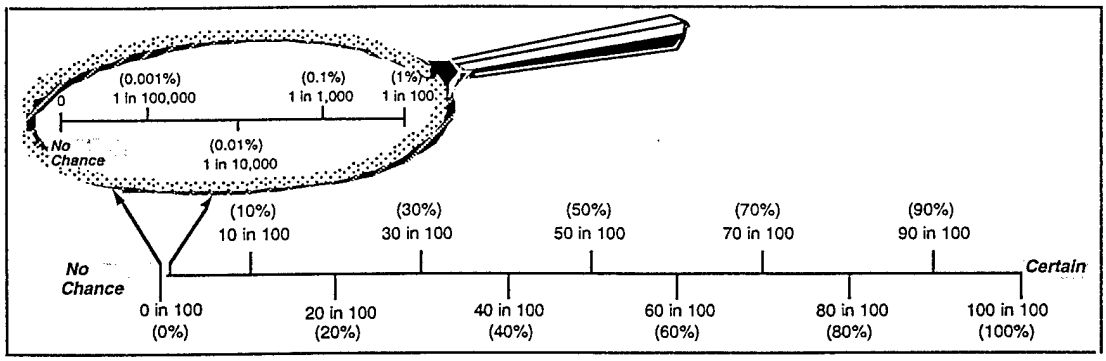
For the first example, we have marked with an "X" the chance of an average person being killed by lightning in the next 10 years. Fortunately, this chance is *very low* so it goes in the magnifying glass.

### Example 1




For practice on the next scale, we would like you to place an "X" for your best guess about the chances of being hit by a meteorite in the next 10 years.

### Practice Question 1

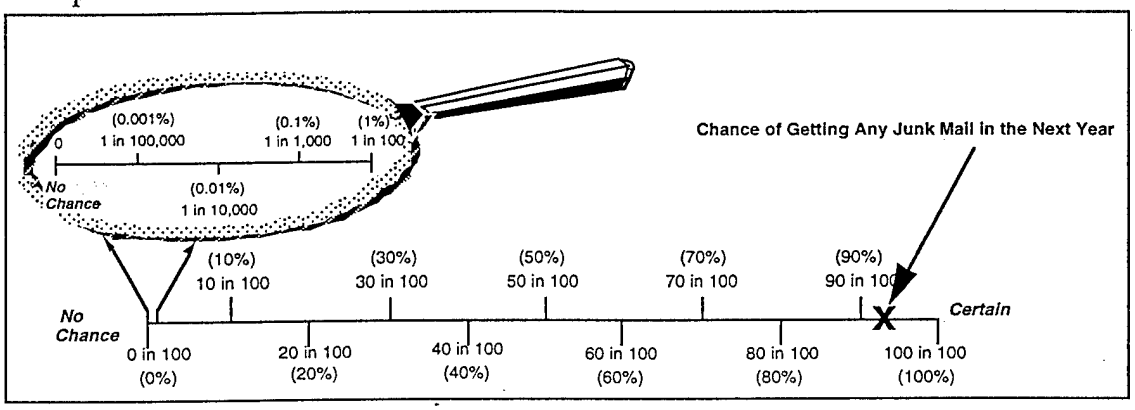



Because the chances are also very low, you should have put your "X" somewhere in the magnifying glass.

*Section 7 continued...*

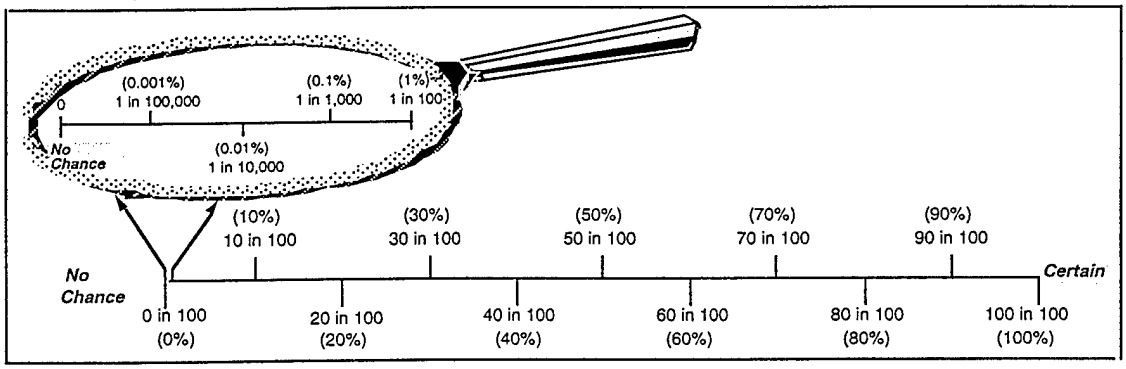
In example 2, we have marked with an "X" the chance of getting junk mail in the next year. Unfortunately, this chance is *very high.*

Example 2




For practice on the next scale, we would like you to place an "X" for your best guess about the chances of stopping at a red light in the next year.

Practice Question 2



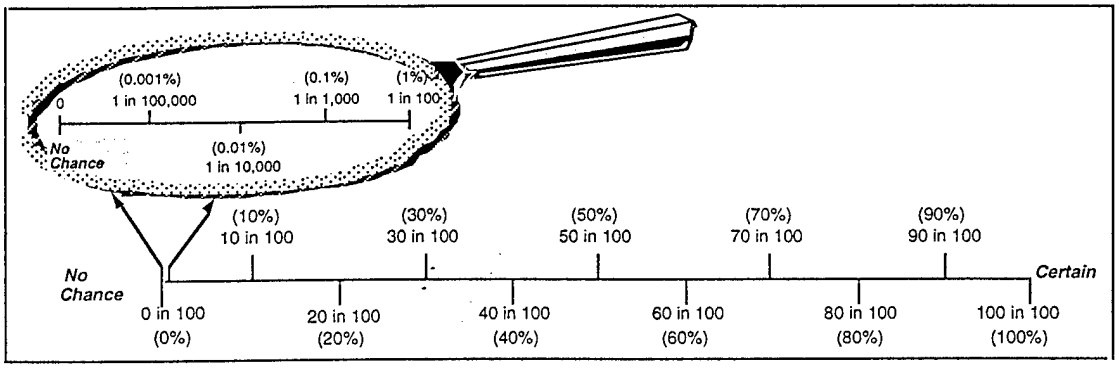

*please turn to next page*

*Section 7 continued...*

**Now for the real questions, we would like you to give your best guess for each of these chances.**

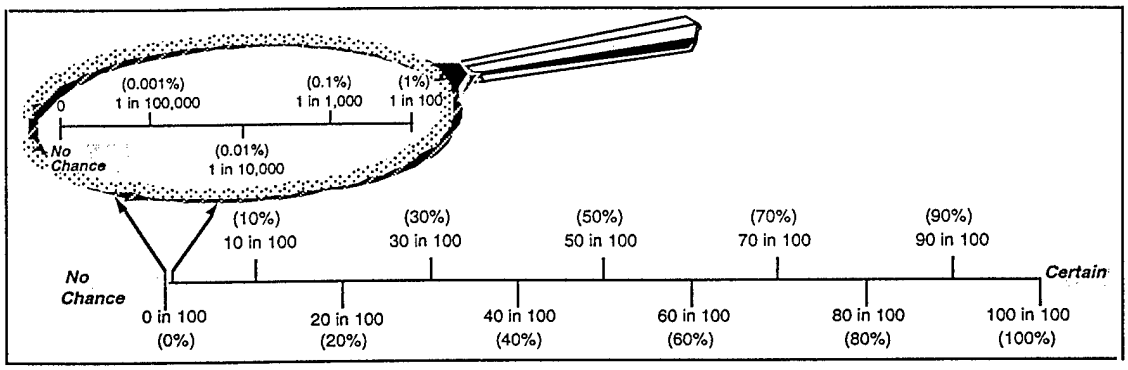
7.01 What is your best guess about the chance that you will be told by a doctor that you <u>have</u> breast cancer sometime in the next 10 years?

Place an "X" on the scale below:




7.02 What is the chance that you will <u>die</u> from breast cancer sometime in the next 10 years?

Place an "X" on the scale below:



*Section 7 continued...*

7.03 How do you think your risk of **getting** breast cancer in the next 10 years compares to that of an average woman your age (Circle **one** response to each question)?

1. My risk is **much** higher
2. My risk is a **little** higher
3. My risk is **about the same**
4. My risk is a **little** lower
5. My risk is **much** lower

7.04 In thinking about all the things that can affect your health, how big of a threat is breast cancer to your health?

1. Very big
2. Big
3. Medium
4. Small
5. Very small
6. Not a threat

Directions for questions 7.05 - 7.08: People have different feelings about various diseases. We would like to understand your feelings about the following. Please rate how much you dread each of the following by placing an "X" anywhere on the dashed line:

7.05 **Being told that you have...**

| | | Don't dread at all | Extremely dreadful |
|---|---|---|---|
| a. | Breast Cancer | [---------- | ----------] |
| b. | Heart Disease | [---------- | ----------] |
| c. | Osteoporosis | [---------- | ----------] |
| d. | Uterine (womb) Cancer | [---------- | ----------] |
| e. | A Blood Clot | [---------- | ----------] |

7.06 **Undergoing treatment for...**

| | | Don't dread at all | Extremely dreadful |
|---|---|---|---|
| a. | Breast Cancer | [---------- | ----------] |
| b. | Heart Disease | [---------- | ----------] |
| c. | Osteoporosis | [---------- | ----------] |
| d. | Uterine (womb) Cancer | [---------- | ----------] |
| e. | A Blood Clot | [---------- | ----------] |

*please turn to next page*

*Section 7 continued...*

7.07 Living with...

| | | Don't dread at all | Extremely dreadful |
|---|---|---|---|
| a. | Breast Cancer | [---------- | ----------] |
| b. | Heart Disease | [---------- | ----------] |
| c. | Osteoporosis | [---------- | ----------] |
| d. | Uterine (womb) Cancer | [---------- | ----------] |
| e. | A Blood Clot | [---------- | ----------] |

7.08 Dying from...

| | | Don't dread at all | Extremely dreadful |
|---|---|---|---|
| a. | Breast Cancer | [---------- | ----------] |
| b. | Heart Disease | [---------- | ----------] |
| c. | Osteoporosis | [---------- | ----------] |
| d. | Uterine (womb) Cancer | [---------- | ----------] |
| e. | A Blood Clot | [---------- | ----------] |

**Directions for questions 7.09 - 7.10: Imagine you have an identical twin. She is not planning to have mammograms. Assume that you know nothing else about her.**

7.09 All things being equal, if your twin got yearly mammograms for the next 10 years, what do you think her chances are of dying from breast cancer (please circle one)?

1. She would have no chance of dying from breast cancer
2. She would have a lower chance of dying from breast cancer
3. There would be no change in her chances of dying from breast cancer
4. She would have a higher chance of dying from breast cancer
5. She will certainly die of breast cancer

*Section 7 continued...*

7.10 What is your best guess about how much your twin's chance of dying from breast cancer would change **with** yearly mammograms?

1. Lower by one-half
2. Lower by one-third
3. Lower by one-fifth to one-tenth
4. No change
5. Higher by one-fifth to one-tenth
6. Higher by one-third
7. Higher by one-half

**Directions for questions 7.11 - 7.17: Your identical twin can do many different things to stay healthy. Please rate the following activities by how much each increases her chances of living for the next 10 years compared to having yearly mammograms for those 10 years by placing an "X" anywhere on the dashed line.**

| | Lowers her chance a lot | Does not change her chance a lot | Increases her chance a lot |
|---|---|---|---|
| 7.11 Compare having just **one** mammogram in 10 years to having yearly mammograms | [---------- | ---------- | ----------] |
| 7.12 Compare **not** smoking cigarettes to having yearly mammograms | [---------- | ---------- | ----------] |
| 7.13 Compare exercising 5 times a week to having yearly mammograms | [---------- | ---------- | ----------] |
| 7.14 Compare eating a low fat diet to having yearly mammograms | [---------- | ---------- | ----------] |
| 7.15 Compare wearing seat belts whenever she rides in a car to having yearly mammograms | [---------- | ---------- | ----------] |
| 7.16 Compare doing a breast self exam every month to having yearly mammograms | [---------- | ---------- | ----------] |
| 7.17 Compare taking estrogen replacement medicine to having yearly mammograms | [---------- | ---------- | ----------] |

*please turn to next page*

## Section 8. How Mammography Relates to You

Directions for question 8.01: **Again, imagine you have an identical twin. (Please circle the appropriate response).**

8.01 How accurate do you think mammograms would be for you compared to your identical twin?

The accuracy of mammograms would be:

1. Much better for me
2. Better for me
3. The same for me
4. Worse for me
5. Much worse for me

Directions for question 8.02: **Your identical twin has an abnormal mammogram. Over the next few weeks, she gets repeat mammograms and a biopsy. It turns out that she doesn't have breast cancer. This kind of abnormal mammogram is called a false alarm. Please circle the appropriate response.**

8.02 Is information about false alarms something you want to factor into your decision about getting a mammogram?

1. Yes 2. No

Directions for question 8.03: **We would like you to fill in the blank in the statement with ONE of the following numbers:**

0 10 50 100 500 1000 10,000 or more

8.03 "I think mammograms are worthwhile even if there were _____ false alarms for each woman's life saved."

*Section 8 continued...*

Directions for questions 8.04 - 8.12: How much do you agree or disagree with the following statements (Circle one number on each line)?

| Response Scale: | Strongly Disagree | Disagree | Neither Agree nor Disagree | Agree | Strongly Agree |
|---|---|---|---|---|---|
| 8.04 I can find the information I need to decide whether to have a mammogram | 1 | 2 | 3 | 4 | 5 |
| 8.05 My personal doctor is the main source of information I need to decide about mammograms | 1 | 2 | 3 | 4 | 5 |
| 8.06 I am confused about whether I should have a mammogram | 1 | 2 | 3 | 4 | 5 |
| 8.07 I have confidence in the recommendations of national expert groups | 1 | 2 | 3 | 4 | 5 |
| 8.08 I am upset when national expert groups disagree about mammograms | 1 | 2 | 3 | 4 | 5 |
| 8.09 If I had all the relevant information, I would know how to use it when making a decision about having a mammogram | 1 | 2 | 3 | 4 | 5 |
| 8.10 The chance of getting breast cancer decreases after menopause | 1 | 2 | 3 | 4 | 5 |
| 8.11 If a woman getting mammograms turns out not to have breast cancer, she may have been harmed by the mammograms | 1 | 2 | 3 | 4 | 5 |
| 8.11a If harms are possible, what are they? _____ | | | | | |
| 8.12 Some types of breast cancer grow so slowly that **even without treatment** they would not affect a woman's health. | 1 | 2 | 3 | 4 | 5 |

*please turn to next page*

## Section 9. How You Feel About Receiving Mammograms

Directions: We would like to understand your experience receiving a mammogram. Please circle the response(s) that best describes your experience.

9.01 The last time you had a mammogram, how would you describe the experience (circle one only).

1. Positive 2. Negative 3. Neutral

9.02 If your response was **negative**, was it due to (circle all that apply):

1. Pain or discomfort during the procedure
2. Difficulty getting an appointment
3. Long waiting time for the exam
4. Procedure skills of the technologist (person performing mammogram)
5. Interpersonal skills of the technologist
6. Other, please describe: _____

9.03 If your response was **positive**, was it due to (circle all that apply):

1. Ease in getting an appointment
2. Getting the procedure done quickly
3. Care taken by the technologist in performing the exam
4. Comfort provided by the technologist in talking you through the exam
5. Opportunity to talk with the radiologist (physician who interprets the exam) about any concerns
6. Other, please describe: _____

9.04 Did the technologist teach you about mammography?

1. Yes 2. No

**If you answered no, please skip to question 9.08, on the next page.**

*Section 9 continued...*

**If you answered yes, please circle the response in questions 9.05-9.07, that best describes your experience.**

| Response Scale: | Strongly Disagree | Disagree | Neither Agree nor Disagree | Agree | Strongly Agree |
|---|---|---|---|---|---|
| 9.05 The information provided to me by the technologist was easy to understand? | 1 | 2 | 3 | 4 | 5 |
| 9.06 The information provided to me by the technologist affected how I feel about getting my next mammogram? | 1 | 2 | 3 | 4 | 5 |
| 9.07 The information provided to you by the technologist made me feel better about getting my next mammogram? | 1 | 2 | 3 | 4 | 5 |

9.08 Are you planning to have a mammogram in the next two years?

1. Definitely yes
2. Probably yes
3. Undecided
4. Probably no
5. Definitely no

**Thank you very much** for taking the time to complete this questionnaire!! We are interested in any thoughts or comments you may have. Please use the space on this page.

# APPENDIX O – HRT DRAFT SURVEY INSTRUMENT

Draft

# NH Women for Health Study

General Instructions: Please answer each question to the best of your ability. If you choose not to answer a particular question, move on to the next question. Please use a No.2 pencil or blue or black pen. Please shade all circles like this: ●
All letters and numbers must be written in capital block style **without touching the sides.**

| 0 | 1 | 2 | 3 | 4 | A | B | C | D | E |
|---|---|---|---|---|---|---|---|---|---|

*The booklet should take about 20 minutes to complete.*

## Section 1: Health History & General Health Habits

1.01 Please fill in all the circles below to show which health events you or your relatives have had:

| HEALTH EVENTS | You | Mother | Father | Aunt | Grand-mother | Grand-father | Sibling | Half-Sibling | Child |
|---|---|---|---|---|---|---|---|---|---|
| Breast Cancer before age 50 | O | O | O | O | O | O | O | O | O |
| Breast Cancer age 50 or older | O | O | O | O | O | O | O | O | O |
| Heart Attack before age 50 | O | O | O | O | O | O | O | O | O |
| Heart attack age 50 or older | O | O | O | O | O | O | O | O | O |
| Other Heart Disease | O | O | O | O | O | O | O | O | O |
| Hip Fracture | O | O | O | O | O | O | O | O | O |
| Other Fracture age 50 or older | O | O | O | O | O | O | O | O | O |
| Endometrial Cancer | O | O | N/A | O | O | N/A | O | O | O |
| Ovarian Cancer | O | O | N/A | O | O | N/A | O | O | O |
| Colon Cancer | O | O | O | O | O | O | O | O | O |
| None of the above | O | O | O | O | O | O | O | O | O |

1.02 Are you currently taking any of the following *(fill-in all that apply):*

- O Blood Pressure Medication
- O Cholesterol Medication
- O Other Heart Medication (such as Beta Blockers)
- O Aspirin, Ibuprofen, Acetominophen, Naproxen
- O Rheumatoid Arthritis Medication (other than listed above)
- O Thyroid Medication (such as Synthroid)
- O Diabetes Medication (such as Insulin)
- O Calcium & Vitamin D Supplement
- O Osteoporosis Medication (such as: Fosamax, Actonel, etc.)
- O Depression Medication (such as: Prozac, Zoloft, Paxil, etc.)
- O Medication for other mental health problems (such as: Haldol, Lithium, Depacpte, etc.)
- O Chemotherapy for Cancer
- O Radiation Therapy for Cancer
- O None of the Above

1.03 Have you ever used corticosteroids (such as Prednisone) for 3 or more months? O Yes O No

Draft

1.04 When did you have your last mammogram? ☐☐ / ☐☐☐☐

M M Y Y Y Y

1.05 Where did you have your last mammogram? _____ _____ ☐☐

Facility Town/city State

1.06 Are you a current or past cigarette smoker (*choose one*)? O Non-smoker O Past smoker O Current smoker

If you are a current or past smoker: a. how many years have/did you smoke in total? ☐☐ Years

b. how many cigarettes do/did you usually smoke per day (*choose one*)?

O 10 or less O 11 - 20 O 21 - 30 O 31 or more

1.07 What best describes your usual exercise routine (*choose one*)?

O I do not exercise regularly

O I exercise 1-2 times a week

O I exercise 3 - 5 times a week

O I exercise more than 5 times a week

1.08 How many alcholic drinks (beer, wine or hard liquor) do you have per week (*choose one*)?

O 0 O 1-4 O 5-8 O 9-12 O 13-16 O 17-20 O 21-25 O 26+

## Section 2: Your Health Today

Directions: By filling in one circle in each group below, please indicate which statements best describe your own health today:

**2.01. Mobility**

O I have no problems in walking about

O I have some problems in walking about

O I am confined to bed

**2.02. Self-Care**

O I have no problems with self-care

O I have some problems washing or dressing myself

O I am unable to wash or dress myself

**2.03. Usual Activities** (e.g. work, study, housework, family or leisure activities)

O I have no problems with performing my usual activities

O I have some problems with performing my usual activities

O I am unable to perform my usual activities

**2.04. Pain/Discomfort**

O I have no pain or discomfort

O I have moderate pain or discomfort

O I have extreme pain or discomfort

**2.05 Anxiety/Depresion**

O I am not anxious or depressed

O I am moderately anxious or depressed

O I am extremely anxious or depressed

Draft

## Section 3: Your Use of Prescription or over the Counter Hormone Therapies (not including birth control pills)

*Instructions:* This section asks you about your use of prescription or over the counter hormone therapies in the following order: 1) if you have ever used prescription hormone therapies; 2) your current use; 3) your previous use; 4) if you have ever used over the counter hormone products; 5) your current use of these products. Please answer each question unless your response leads you to instructions to skip to another section.

3.01 Please fill in the circles below to show what type of prescription hormone therapies you have ever taken for three or more months in a row.

| ○ I never used prescription hormone therapies | Estrogen Only | | Progestin & Estrogen Combined | Progestin Only | Other |
|---|---|---|---|---|---|
| | Patch Form | Pill Form | Pill Form Only | Pill Form Only | Pill or Cream |
| **Please go to the top of page 6** | ○ Alora<br>○ Climera<br>○ Estraderm<br>○ Fempatch<br>○ Vivelle | ○ Premarin<br>○ Estrace<br>○ Estratab<br>○ Menest<br>○ Ortho-Est<br>○ DES | ○ Prempro<br>○ Premphase | ○ Provera<br>○ Amen<br>○ Aygestin<br>○ Cycrin<br>○ Prometrium | ○ Evista (Raloxifene)<br>○ Nolvadex (Tamoxifen)<br>○ Testosterone<br>○ DHEA<br>○ Estratest<br>○ Pregnenolene |

3.02 What is the total length of time you have used any type of prescription hormone therapies? (*choose one*)

| | | | |
|---|---|---|---|
| ○ 3 - 6 months | ○ 3 years | ○ 6 years | ○ 9 years |
| ○ 1 year | ○ 4 years | ○ 7 years | ○ 10 - 19 years |
| ○ 2 years | ○ 5 years | ○ 8 years | ○ 20 years or more |

3.03 What are the main reason(s) you first started using prescription hormone therapies? (*fill in all that apply*):

| | | |
|---|---|---|
| ○ Hysterectomy | ○ Hot Flashes | ○ Depression & mood swings |
| ○ Night Sweats | ○ Vaginal dryness | ○ Concern about heart disease |
| ○ Irregular Bleeding | ○ Prescribed by physician | ○ Other - please specify _____ |
| ○ Menstrual pain/cramps | ○ Concern about osteoporosis (brittle bones) or fractures | _____ |

3.04 Please fill in the circles below to show what type of prescription hormone therapies you have used most recently.

| Estrogen Only | | Progestin & Estrogen Combined | Progestin Only | Other |
|---|---|---|---|---|
| Patch Form | Pill Form | Pill Form Only | Pill Form Only | Pill or Cream |
| ○ Alora<br>○ Climera<br>○ Estraderm<br>○ Fempatch<br>○ Vivelle<br>○ Other | ○ Premarin<br>○ Estrace<br>○ Estratab<br>○ Menest<br>○ Ortho-Est<br>○ DES<br>○ Other | ○ Prempro<br>○ Premphase<br>○ Other | ○ Provera<br>○ Amen<br>○ Aygestin<br>○ Cycrin<br>○ Prometrium<br>○ Other | ○ Evista (Raloxifene)<br>○ Nolvadex (Tamoxifen)<br>○ Testosterone<br>○ DHEA<br>○ Estratest<br>○ Pregnenolene<br>○ Other |

Draft

3.05 Did you use this type of prescription hormone therapy *in the past 6 months*?

O Yes

3.05a For about how long have you used this prescription hormone therapy (*choose one*)?

- O Less than 3 months
- O 3-6 months
- O 1 year
- O 2 years
- O 3 years
- O 4 years
- O 5 years
- O 6 years
- O 7 years
- O 8 years
- O 9 years
- O 10 - 19 years
- O 20 years or more

3.05b How old were you when you started taking it? ☐☐ Age

3.05c Are you still using it?

- O Yes
- O No ⟶ **If no**, how many months ago did you stop taking it? ☐☐ months

3.05d During the past 6 months, have you experienced any of the following side effects as a result of prescription hormone therapies (*fill-in all that apply*)?

- O no side effects
- O irregular menstrual-like bleeding
- O regular menstrual-like bleeding
- O breast tenderness (or discomfort)
- O weight gain of 5 or more pounds
- O headache or flu-like symptoms
- O depression or mood swings
- O nausea/stomach pain
- O bloating/fluid retention
- O blood clots in legs and/or lungs
- O leg cramps
- O hot flashes
- O other - please specify _____

3.05e Overall, how bothersome have these side effects been?

- O no side effects
- O not at all bothersome
- O a little bothersome
- O very bothersome
- O extremely bothersome

O No

3.05g Approximately how long has it been since you last used this prescription hormon therapy (*choose one*)?

- O 1 year
- O 2 years
- O 3 years
- O 4 years
- O 5 years
- O 6 years
- O 7 years
- O 8 years
- O 9 years
- O 10 - 19 years
- O 20 years or more

3.05h How old were you when you started taking it? ☐☐ Years

3.05i How old were you when you stopped taking it? ☐☐ Years

3.05j What were the main reasons(s) you stopped taking it (*fill-in all that apply*)?

- O irregular menstrual-like bleeding
- O regular menstrual-like bleeding
- O leg cramps
- O hot flashes
- O other side effects _____
- O I felt I didn't need it
- O cost of medication was not worth it
- O physician recommended that I stop
- O concern about breast cancer
- O concern about endometrial (uterine) cancer
- O other - Please specify:_____

3.06 Did you take any other type of prescription hormone therapies, before the one you just told us about, for 3 or more months in a row?

O Yes ↓ O No → **Please go to the top of page 6**

3.07 Please report the type of prescription hormone therapies you used before the one you just told us about:

| Estrogen Only | | Progestin & Estrogen Combined | Progestin Only | Other |
|---|---|---|---|---|
| Patch Form | Pill Form | Pill Form Only | Pill Form Only | Pill or Cream |
| O Alora | O Premarin | O Prempro | O Provera | O Evista (Raloxifene) |
| O Climera | O Estrace | O Premphase | O Amen | O Nolvadex (Tamoxifen) |
| O Estraderm | O Estratab | O Other | O Aygestin | O Testosterone |
| O Fempatch | O Menest | | O Cycrin | O DHEA |
| O Vivelle | O Ortho-Est | | O Prometrium | O Estratest |
| O Other | O DES | | O Other | O Pregnenolene |
| | O Other | | | O Other |

3.08 How old were you when you started taking it? ☐☐ Years

3.09 How old were you when you stopped taking it? ☐☐ Years

3.10 Did you use another type of prescription hormone therapy, before the one you described above, for 3 or more months in a row?

O Yes ↓ O No → **Please go to the top of page 6**

3.11 Please report the type of prescription hormonal therapy you used before the one you just told us about:

| Estrogen Only | | Progestin & Estrogen Combined | Progestin Only | Other |
|---|---|---|---|---|
| Patch Form | Pill Form | Pill Form Only | Pill Form Only | Pill or Cream |
| O Alora | O Premarin | O Prempro | O Provera | O Evista (Raloxifene) |
| O Climera | O Estrace | O Premphase | O Amen | O Nolvadex (Tamoxifen) |
| O Estraderm | O Estratab | O Other | O Aygestin | O Testosterone |
| O Fempatch | O Menest | | O Cycrin | O DHEA |
| O Vivelle | O Ortho-Est | | O Prometrium | O Estratest |
| O Other | O DES | | O Other | O Pregnenolene |
| | O Other | | | O Other |

3.12 How old were you when you started taking it? ☐☐ Years

3.13 How old were you when you stopped taking it? ☐☐ Years

3.14 During the past five years, have you taken any other type of perscription hormonal therapies for 3 or more months in a row that you have not already told us about? O Yes O No

Please continue on the top of the next page.

Draft

3.15 Did a doctor ever prescribe prescription hormone therapies for you that you decided not to use? O Yes O No

3.16 Please indicate the main reason(s) you decided not to use prescription hormonal therapies (*fill in all that apply*):

- O It never occurred to me
- O I take dietary hormone supplements
- O I am still having menstrual periods
- O I am concerned about potential side effects
- O I feel I don't need it
- O I think medication might be harmful
- O I don't like to use medications
- O Cost of medication is not worth it
- O Insurance does not cover it
- O I want to experience menopause naturally
- O I am concerned about breast cancer
- O I am concerned about endometrial (uterine) cancer
- O I have never heard of it

3.17 Have you ever used any kind of hormone supplements (dietary or over the counter) at least weekly for 3 or more months in a row (such as: teas, other liquid supplements, foods)?

| O Yes ↓ | O No ↓ |
|---|---|
| 3.17a What type of dietary hormone supplement(s) have you used (*fill-in all that apply*)?<br>O Black Cohosh (such as Remifemin) - in capsule, tincture, tablet or tea form<br>O Soy or soy supplement(s)<br>O Homeopathic remedies (pellets dissolved under the tongue)<br>O Glandulars/Protomorphogens (ovarian hormone extracts from organically grown animals)<br>O Herbal remedies (such as: Fem-H, PMS Herbal, Herbal F, Spectra Ostaderm)<br>O Other teas, tinctures, food supplements, tablets, or capsules<br>O Other over the counter hormone supplement (please specify): _____<br><br>3.17b Please estimate the total length of time you have used any type of over the counter hormone supplements:<br>O 3-6 months O 6 years<br>O 1 year O 7 years<br>O 2 years O 8 years<br>O 3 years O 9 years<br>O 4 years O 10-19 years<br>O 5 years O 20 years or more | 3.17c What are the main reason(s) you decided not to use dietary or over the counter hormone supplements (*fill-in all that apply*):<br>O It never occurred to me<br>O I take prescription hormonal therapy<br>O I am still having menstrual periods<br>O I am concerned about potential side effects<br>O I feel I don't need it<br>O I think it might be harmful<br>O I don't like to use dietary supplements<br>O Cost of supplements is not worth it<br>O Insurance does not cover it<br>O I want to experience menopause naturally<br>O I am concerned about breast cancer<br>O I am concerned about endometrial (uterine) cancer<br>O I have never heard of it<br>O Other _____<br><br>**Please Go to Section 4, next page.** |

Draft

3.18 Did you use dietary or over the counter hormone supplements *in the past six months*?

| O Yes ↓ | O No ↓ |
|---|---|
| 3.19 If yes, what supplements did you use? (*fill-in all that apply*)<br>O Black Cohosh (Remifemin) - in capsule, tincture, tablet or tea form<br>O Soy supplements<br>O Homeopathic remedies (pellets dissolved under the tongue)<br>O Glandulars/Protomorphogens (ovarian hormone extracts from organically grown animals)<br>O Herbal remidies (Fem-H, PMS Herbal, Herbal F, Spectra Ostaderm)<br>O Other teas, tinctures, food supplements, tablets, or capsules<br>O Other over the counter hormone supplement (please specify): _____<br>3.20 How old were you when you started taking it? ☐☐ years<br>3.21 Please write the total number of supplements you usually take per day or per week (for example, if you take a tea and a soy supplement each day you should write "02" in the per day box):<br>☐☐ per day or ☐☐ per week<br>3.22 During the past 6 months have you experienced any of the following side effects as a result of the hormone supplements you have taken? (fill-in all that apply):<br>O irregular menstrual-like bleeding? O nausea/stomach pain<br>O regular menstrual-like bleeding? O depression or mood swings<br>O breast tenderness(discomfort) O bloating/fluid retention<br>O weight gain of 5 pounds or more O blod clots in legs and-or lungs<br>O headache or flu-like symptoms O hot flashes | 3.23 If no, how many months ago did you stop taking it?<br>☐☐<br>Months<br>**Please go to section 4, below.** |

## Section 4 Women's Health Questionnaire

Directions: Please indicate how you are feeling now, or how you have been feeling in the last few days, by filling in the correct circle to answer each of the following statements:

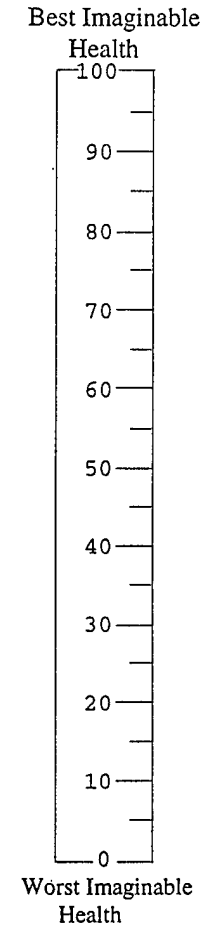
| | Yes definitely | Yes sometimes | No not much | No not at all |
|---|---|---|---|---|
| 4.01 I wake early, then sleep badly for the rest of the night. | O | O | O | O |
| 4.02 I get very frightened or panic feelings for apparently no reason at all. | O | O | O | O |
| 4.03 I feel miserable and sad. | O | O | O | O |
| 4.04 I feel anxious when I go out of the house on my own. | O | O | O | O |

| | Yes definitely | Yes sometimes | No not much | No not at all |
|---|---|---|---|---|
| 4.05 I have lost interest in things. | O | O | O | O |
| 4.06 I get palpitations or a sensation of "butterflies" in my stomach or chest. | O | O | O | O |
| 4.07 I still enjoy the things I used to. | O | O | O | O |
| 4.08 I feel life is not worth living. | O | O | O | O |
| 4.09 I feel tense or "wound up". | O | O | O | O |
| 4.10 I have a good appetite. | O | O | O | O |
| 4.11 I am restless and can't keep still. | O | O | O | O |
| 4.12 I am more irritable than usual. | O | O | O | O |
| 4.13 I worry about growing old. | O | O | O | O |
| 4.14 I have headaches. | O | O | O | O |
| 4.15 I feel more tired than usual. | O | O | O | O |
| 4.16. I have dizzy spells. | O | O | O | O |
| 4.17. My breasts feel tender or uncomfortable | O | O | O | O |
| 4.18. I suffer from backaches/or pains in my limbs. | O | O | O | O |
| 4.19. I have hot flashes. | O | O | O | O |
| 4.20. I am more clumsy than usual. | O | O | O | O |
| 4.21. I feel rather lively and excitable. | O | O | O | O |
| 4.22. I have abdominal cramps or discomfort. | O | O | O | O |
| 4.23. I feel sick or nauseous. | O | O | O | O |
| 4.24. I have lost interest in sexual activity. | O | O | O | O |
| 4.25. I have feelings of well-being. | O | O | O | O |
| 4.26. I have heavy (menstrual) periods. | O | O | O | O |
| 4.27. I suffer from night sweats. | O | O | O | O |
| 4.28. My stomach feels bloated. | O | O | O | O |
| 4.29. I have difficulty in getting off to sleep. | O | O | O | O |
| 4.30. I often notice pins and needles in my hands and feet. | O | O | O | O |
| 4.31 I am satisfied with my current sexual relationship. (please omit if not sexually active) | O | O | O | O |

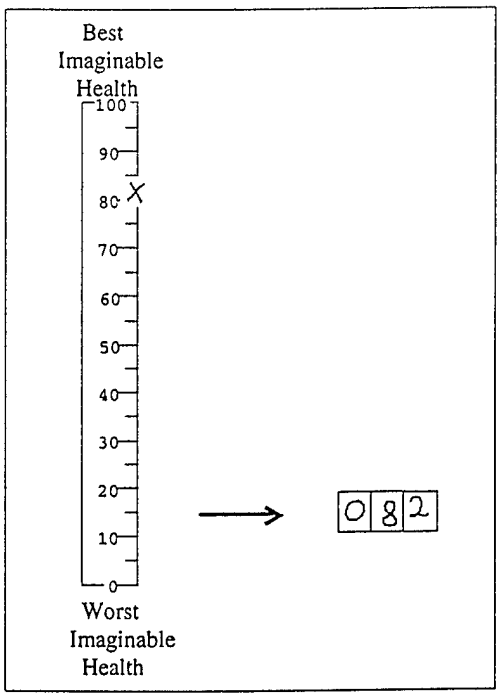
| | Yes definitely | Yes sometimes | No not much | No not at all |
|---|---|---|---|---|
| 4.32. I feel physically attractive. | O | O | O | O |
| 4.33. I have difficulty in concentrating. | O | O | O | O |
| 4.34. As a result of vaginal dryness, sexual intercourse has become uncomfortable. (please omit if not sexually active) | O | O | O | O |
| 4.35. I need to pass urine/water more frequently than usual. | O | O | O | O |
| 4.36 My memory is poor. | O | O | O | O |

## Section 5 Global Health Question

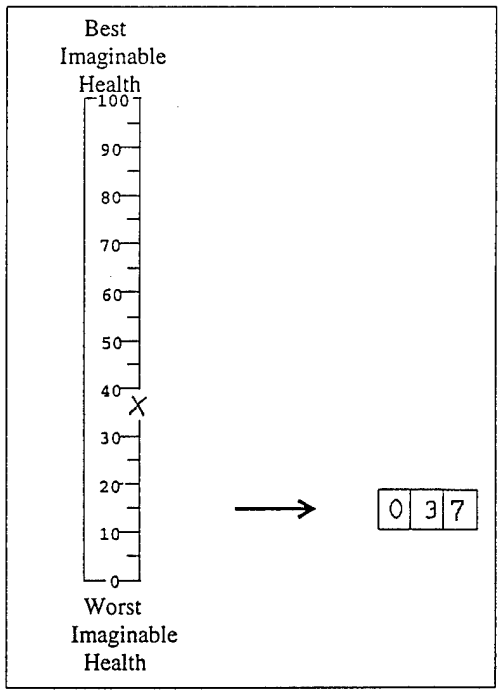
5.01 The scale at the right ranges from 100 (best imaginable health) to 0 (worst imaginable health). Please rate how you feel about your health <u>today</u> by placing an "X" on the scale to the right. Two examples using this scale are shown below.




EXAMPLE 1




EXAMPLE 2




As shown in the examples above, please write the value you marked with an "X" in these boxes. → | | | |

Draft

## Section 6 Your General Health

Directions: Please indicate which statement best describes your own health state over the past 4 weeks by filling in one circle below.

6.01 In general would you say your health is:

O Excellent O Very good O Good O Fair O Poor

6.02 The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much (*Fill in one circle on each line*)

| | Yes, Limited a lot | Yes, limited a little | No, not limited at all |
|---|---|---|---|
| a. **Vigorous activities** such as running, lifting heavy objects, or participating in strenuous sports | O | O | O |
| b. **Moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling or playing golf | O | O | O |
| c. Lifting or carrying groceries | O | O | O |
| d. Climbing **several** flights of stairs | O | O | O |
| e. Climbing **one** flight of stairs | O | O | O |
| f. Bending, kneeling, or stooping | O | O | O |
| g. Walking **more than a mile** | O | O | O |
| h. Walking **several blocks** | O | O | O |
| i. Walking **one block** | O | O | O |
| j. Bathing or dressing yourself | O | O | O |

6.03 During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

| | (*fill in one circle on each line*) YES | NO |
|---|---|---|
| a. Cut down the **amount of time** you spent on work or other activities | O | O |
| b. **Accomplished less** than you would like | O | O |
| c. Were limited in the **kind** of work or other activities | O | O |
| d. Had **difficulty** performing the work or other activities (for example, it took extra effort) | O | O |

6.04. During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

*(fill in one circle on each line)*

| | YES | NO |
|---|---|---|
| a. Cut down the **amount of time** you spent on work or other activities | O | O |
| b. **Accomplished less** than you would like | O | O |
| c. Didn't do work or other activities as **carefully** as usual | O | O |

6.05 During the past 4 weeks, to what extent has your physical health or emotional problems intefered with your normal social activities with family, friends, neighbors, or groups? *(fill in one circle)*

O Not at all O Slightly O Moderatly O Quite a bit O Extremely

6.06 How much bodily pain have you had during the past 4 weeks? *(fill in one circle)*

O None O Very mild O Mild O Moderate O Severe O Very severe

6.07 During the past 4 weeks, how much did pain intefere with your normal work (including both work outside the home and housework) *(fill in one circle)*.

O Not at all O Slightly O Moderatly O Quite a bit O Extremely

6.08 These questions are about how you feel and how things have been with you during the past four weeks. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the past 4 weeks. *(fill in one circle on each line)*

| | All of the time | Most of the time | A good bit of the time | Some of the time | A little of the time | None of the time |
|---|---|---|---|---|---|---|
| 1 Did you feel full of pep? | O | O | O | O | O | O |
| 2 Have you been a very nervous person? | O | O | O | O | O | O |
| 3 Have you felt so down in the dumps that nothing could cheer you up? | O | O | O | O | O | O |
| 4 Have you felt calm and peaceful? | O | O | O | O | O | O |
| 5 do you have a lot of energy? | O | O | O | O | O | O |
| 6 Have you felt downhearted and blue? | O | O | O | O | O | O |
| 7 Did you feel worn out? | O | O | O | O | O | O |
| 8 Have you been a happy person | O | O | O | O | O | O |
| 9 Did you feel tired? | O | O | O | O | O | O |

Draft

6.09. During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)? (*fill in one circle*)

O All of the time O Most of the time O Some of the time

O A little of the time O None of the time

6.10. How TRUE or FALSE is each of the following statements for you. (*fill in one circle on each line*)

| | Definitely True | Mostly True | Don't know | Mostly False | No not at all |
|---|---|---|---|---|---|
| a. I seem to get sick a little easier than other people | O | O | O | O | O |
| b. I am as healthy as anybody I know | O | O | O | O | O |
| c. I expect my health to get worse | O | O | O | O | O |
| d. My health is excellent | O | O | O | O | O |

## Section 7 Your Future Health

7.01. Compared to an average woman your age, please give your best guess about your chance of being told by a doctor in the next 10 years that you have (*fill in one circle on each line*):

| | Much lower chance | Little lower chance | Same chance | Little higher chance | Much higher chance | I Already have this |
|---|---|---|---|---|---|---|
| a. Breast cancer | O | O | O | O | O | O |
| b. Heart disease | O | O | O | O | O | O |
| c. osteoporosis (brittle bones) | O | O | O | O | O | O |

We want to be sure we understand your answers to this survey. We would like to contact you if we have questions. Please include your telephone number here if we may contact you with brief questions:

8 0 2 - [ ] [ ] [ ] - [ ] [ ] [ ] [ ]

***Thank you again for participating in our study***